CORRECTED VERSION*

AD

PCT - WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, C12P 19/34, C07H 21/02, 21/04

A1

(11) International Publication Number: WO 97/46710

(43) International Publication Date: 11 December 1997 (11.12.97)

(21) International Application Number: PCT/US97/09774

(22) International Filing Date: 5 June 1997 (05.06.97)

(30) Priority Data:

08/658,578	5 June 1996 (05.06.96)	US
08/846,111	25 April 1997 (25.04.97)	US

(81) Designated States: AU, CA, CN, JP, KR, NZ, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant: LUDWIG INSTITUTE FOR CANCER RESEARCH [CH/US]; 1345 Avenue of the Americas, New York, NY 10105 (US).

(72) Inventors: LURQUIN, Christophe; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE). BRASSEUR, Francis; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE). BOON, Thierry; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE).

(74) Agent: HANSON, Norman, D.; Felfe & Lynch, 805 Third Avenue, New York, NY 10022-7513 (US).

(54) Title: ISOLATED NUCLEIC ACID MOLECULES WHICH ARE MEMBERS OF THE MAGE-B FAMILY AND USES THEREOF

(57) Abstract

The invention relates to members of the MAGE-B family of nucleic acid molecules. These molecules differ from the previously described MAGE nucleic acid molecules in that members of the MAGE-Xp family do not hybridize to the previously identified MAGE sequences. Further, the members of the MAGE-B family are found on the Xp arm of the X chromosome rather than on the Xq chromosome, as was the case with the previously identified MAGE genes.

* (Referred to in PCT Gazette No. 08/1998, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12P 19/34, C07H 21/02, 21/04 — A1

(11) International Publication Number: **WO 97/46710**

(43) International Publication Date: 11 December 1997 (11.12.97)

(21) International Application Number: PCT/US97/09774

(22) International Filing Date: 5 June 1997 (05.06.97)

(30) Priority Data:

08/658,578	5 June 1996 (05.06.96)	US
Not furnished	25 April 1997 (25.04.97)	US

(81) Designated States: AU, CA, CN, JP, KR, NZ, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant: LUDWIG INSTITUTE FOR CANCER RESEARCH [CH/US]; 1345 Avenue of the Americas, New York, NY 10105 (US).

(72) Inventors: LURQUIN, Christophe; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE). BRASSEUR, Francis; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE). BOON, Thierry; Avenue Hippocrate 74, UCL 7459, B-1200 Brussels (BE).

(74) Agent: HANSON, Norman, D.; Felfe & Lynch, 805 Third Avenue, New York, NY 10022-7513 (US).

(54) Title: ISOLATED NUCLEIC ACID MOLECULES WHICH ARE MEMBERS OF THE MAGE-B FAMILY AND USES THEREOF

(57) Abstract

The invention relates to members of the MAGE-B family of nucleic acid molecules. These molecules differ from the previously described MAGE nucleic acid molecules in that members of the MAGE-Xp family do not hybridize to the previously identified MAGE sequences. Further, the members of the MAGE-B family are found on the Xp arm of the X chromosome rather than on the Xq chromosome, as was the case with the previously identified MAGE genes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

ISOLATED NUCLEIC ACID MOLECULES WHICH ARE MEMBERS OF THE MAGE-B FAMILY AND USES THEREOF

RELATED APPLICATIONS

This application is a continuation-in-part of Serial No. 08/658,578, filed June 5, 1996, which is a continuation-in-part of Serial No. 08/403,388, filed March 14, 1995, both of which are incorporated by reference.

FIELD OF THE INVENTION

This invention relates to a nucleic acid molecule which codes for a tumor rejection antigen precursor. More particularly, the invention concerns genes, whose tumor rejection antigen precursor is processed, inter alia, into at least one tumor rejection antigen. The tumor rejection antigen precursors in question do not appear to be closely related to other known tumor rejection antigen precursor coding sequences, and were isolated from the Xp region of human X chromosomes, in contrast to the genes to which they are most closely related, which were found on the Xq region. These newly isolated genes are members of the MAGE-B family, while those in the Xq region are now considered to be members of the MAGE-A family.

BACKGROUND AND PRIOR ART

The process by which the mammalian immune system recognizes and reacts to foreign or alien materials is a complex one. An important facet of the system is the T lymphocyte, or "T cell" response. This response requires that T cells recognize and interact with complexes of cell surface molecules, referred to as human leukocyte antigens ("HLA"), or major histocompatibility complexes ("MHCs"), and peptides. The peptides are derived from larger molecules which are processed by the cells which also present the HLA/MHC molecule. See in this regard Male et al., Advanced Immunology (J.P. Lipincott Company, 1987), especially chapters 6-10. The interaction of T cells and HLA/peptide complexes is restricted, requiring a T cell specific for a particular combination of an HLA molecule and a peptide. If

a specific T cell is not present, there is no T cell response even if its partner complex is present. Similarly, there is no response if the specific complex is absent, but the T cell is present. This mechanism is involved in the immune system's response to foreign materials, in autoimmune pathologies, and in responses to cellular abnormalities. Much work has focused on the mechanisms by which proteins are processed into the HLA binding peptides. See, in this regard, Barinaga, Science 257: 880 (1992); Fremont et al., Science 257: 919 (1992); Matsumura et al., Science 257: 927 (1992); Latron et al., Science 257: 964 (1992).

The mechanism by which T cells recognize cellular abnormalities has also been implicated in cancer. For example, in PCT application PCT/US92/04354, filed May 22, 1992, published on November 26, 1992, and incorporated by reference, a family of genes is disclosed, which are processed into peptides which, in turn, are expressed on cell surfaces, which can lead to lysis of the tumor cells by specific CTLs cytolytic T lymphocytes, or "CTLs" hereafter. The genes are said to code for "tumor rejection antigen precursors" or "TRAP" molecules, and the peptides derived therefrom are referred to as "tumor rejection antigens" or "TRAs". See Traversari et al., Immunogenetics 35: 145 (1992); van der Bruggen et al., Science 254: 1643 (1991), for further information on this family of genes. Also, see U.S. patent application Serial Number 807,043, filed December 12, 1991, now U.S. Patent No. 5,342,774, incorporated by reference in its entirety. The "MAGE" family of tumor rejection antigen precursors is disclosed in this patent.

In U.S. patent application Serial Number 938,334, now U.S. Patent No. 5,405,940, April 15, 1995, the disclosure of which is incorporated by reference, it is explained that the MAGE-1 gene codes for a tumor rejection antigen precursor which is processed to nonapeptides which are presented by the HLA-A1 molecule. The nonapeptides which bind to HLA-A1 follow a "rule" for binding in that a motif is satisfied.

In this regard, see e.g. PCT/US93/07421; Falk et al., Nature 351: 290-296 (1991); Engelhard, Ann Rev. Immunol. 12: 181-207 (1994); Ruppert et al., Cell 74: 929-937 (1993); Rötzschke et al., Nature 348: 252-254 (1990); Bjorkman et al., Nature 329: 512-518 (1987); Traversari et al., J. Exp. Med. 176: 1453-1457 (1992). The references teach that given the known specificity of particular peptides for particular HLA molecules, one should expect a particular peptide to bind to one HLA molecule, but not to others. This is important, because different individuals possess different HLA phenotypes. As a result, while identification of a particular peptide as being a partner for a specific HLA molecule has diagnostic and therapeutic ramifications, these are only relevant for individuals with that particular HLA phenotype. There is a need for further work in the area, because cellular abnormalities are not restricted to one particular HLA phenotype, and targeted therapy requires some knowledge of the phenotype of the abnormal cells at issue.

In U.S. Patent Application Serial Number 008,446, filed January 22, 1993 and incorporated by reference, the fact that the MAGE-1 expression product is processed to a second TRA is disclosed. This second TRA is presented by HLA-Cw*1601 molecules. The disclosure shows that a given TRAP can yield a plurality of TRAs, each of which will satisfy a motif rule for binding to an MHC molecule.

In U.S. Patent Application Serial Number 994,928, filed December 22, 1992, and incorporated by reference herein teaches that tyrosinase, a molecule which is produced by some normal cells (e.g., melanocytes), is processed in tumor cells to yield peptides presented by HLA-A2 molecules.

In U.S. patent application Serial No. 08/032,978, filed March 18, 1993, and incorporated by reference in its entirety, a second TRA, not derived from tyrosinase is taught to be presented by HLA-A2 molecules. The TRA is derived from a TRAP, but is coded for by a non-MAGE gene. This disclosure shows that a particular HLA molecule may present TRAs derived from different sources.

In U.S. patent application Serial No.08/079,110, filed June 17, 1993 and incorporated by reference herein, an unrelated tumor rejection antigen precursor, the so-called "BAGE" precursor is described. The BAGE precursor is not related to the MAGE family.

In U.S. patent applications Serial No. 08/096,039 and Serial No. 08/250,162, both of which are incorporated by reference, non-related TRAP precursor GAGE is also disclosed.

The work which is presented by the papers, patent, and patent applications cited supra deal, in large part, with the MAGE family of genes, and the unrelated BAGE, GAGE and DAGE genes, showing that there are different, additional tumor rejection antigen precursors expressed by cells.

It has now been found that there is yet another family of tumor rejection antigen precursor genes. These nucleic acid molecules show homology to the MAGE family of genes, but this homology is insufficient to identify the members of the MAGE-B family by hybridization with the members of the MAGE-A family, as set forth in, e.g., PCT Application PCT/US92/04354 and U.S. Patent No. 5,342,774, under the conditions of stringency set forth therein. Further, the isolated nucleic acid molecules of the invention were all found on the Xp arm of the X chromosome, as contrasted to the previously identified members of the MAGE-A family, all of which were found on the Xq arm. Thus, the invention relates to isolated nucleic acid molecules which encode for MAGE-B tumor rejection antigen precursors and the uses thereof.

The invention is explained in further detail in the disclosure which follows.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Example 1

The cosmids D5 and 4965 have been described by Muscatelli, et al., Nature 372: 672-676 (1994), as well as in Muscatelli, et al., Proc. Natl. Acad. Sci. USA 92: 4987-4991 (1995) the disclosures of which are incorporated by

reference. These cosmids contain portions of the Xp arm of the X-chromosome. The cosmids were digested, using restriction endonucleases EcoRI, BamHI, Hind III, and PstI. Once digested, the DNA was transferred, to a nylon membrane, following agarose electrophoretic migration in an agarose gel.

Following this, a probe, based upon SEQ ID NO: 1, i.e., the sequence for Xp1, was used in hybridization experiments. The probe was approximately 0.45 kilobases in length, and contains 41 base pairs of the first exon (73 base pairs total), the complete second exon, and 299 base pairs of the third (1603 base pairs total). The sequence for what is referred to herein as "MAGE-B1" and is referred to elsewhere as "Xp" may be found in Muscatelli, et al., Proc. Natl. Acad. Sci. USA supra. Further the sequence is found in the EMBL sequence data bank reference to accession number emb X82539, available no later than February 7, 1995.

In order to prepare the 0.4 kb probe, the following primers, i.e., SEQ ID NO: 11 and SEQ ID NO: 12 were used, in PCR, on B1 cDNA:

5'-GTGGTGTCCAGCAGTGTCTC-3'

5'-GTCAGATTCGGTACATGACACAG-3

Specifically, the DNA was denatured with NaOH and neutralized in the gel before transfer to a nylon membrane using 20xSSC (SSC=0.15M NaCl, 0.015M sodium citrate, pH 7). Following transfer, the membranes were rinsed for 5 minutes in 6xSSC at room temperature, baked for one hour at 80°C, and pretreated for 4 hours in 6xSSC, 10xDenhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% BSA), at 65°C.

The membrane was then hybridized in 3.5xSSC, 1xDenhardt's Solution, 25 mm sodium phosphate buffer (pH 7.0), 0.5% SDS, 2mM EDTA and $3x10^6$ cpm/ml α ^{32}P-CTP radiolabelled probe. Hybridization was performed for 18 hours at 65°C. The membrane was then washed at 65°C, four times, for one hour each time in 2xSSC, 0.5% SDS, 1xDenhardt's solution; once for 30 minutes at 0.2xSSC, 0.1%

SDS; and once for 30 minutes in 0.1xSSC, 0.1% SDS. The membranes were autoradiographed using Kodak X-ARS film, and Kodak X-Omatic fine intensifying screens.

Following the hybridization, several signals of differing intensity were observed. Of these, three EcoRI fragments from cosmid 4965, which were 1.5, 2.2, and 2.5 kilobases in length were isolated, and cloned into vector pTZ19R for sequencing. Partial sequencing showed that each fragment contained a sequence homologous to the third exon of B1. Homology of the three sequences, relative to B1, was 75%, 60%, and 80%, for genes referred to hereafter as MAGE-B2, MAGE-B3, and MAGE-B4. These are presented in SEQ ID NOS: 2, 3 and 4, respectively.

The foregoing disclosure, places many tools of extreme value in the hands of the skilled artisan. To begin, the examples identify isolated nucleic acid molecules which code for MAGE-B tumor rejection antigen precursors as well as the nucleic acid molecules complementary thereto. It is known that DNA exists in double stranded form, and that each of the two strands is complementary to the other. Nucleic acid hybridization technology has developed to the point where, given a strand of DNA, the skilled artisan can isolate its complement, or synthesize it. The invention includes, inter alia, the phenomenon of double strandedness to permit the artisan to identify the X chromosome, especially the Xp element, as well as defects in the chromosome.

Such assays can be carried out by one of ordinary skill in the art, using standard methodologies. For example, using the well known polymerase chain reaction (PCR), one uses the following primers:

For identifying B2:

5'-TAAAAAAGGTGCCAAGAGCCAC-3' (SEQ ID NO: 5);

5'-TGAGGCCCTCAGAGGCTTTC-3' (SEQ ID NO: 6).

For identifying B3:

5'-AGTCTGCTGGTAGGTCACGTA-3' (SEQ ID NO: 7);

5'-TCAGGAACTGCACCAACATATTT-3' (SEQ ID NO: 8).

For identifying B4:

5'-AGGGATACTGCCTCCAGCTC-3' (SEQ ID NO: 9);

5'-CAGGAACTGCACTAACATCTTC-3' (SEQ ID NO: 10).

Example 2, which follows, shows one way this can be done.

EXAMPLE 2

The primers of SEQ ID NO: 5 and SEQ ID NO: 6 were used, for example, to determine whether or not MAGE-B2 was expressed in tumors.

Total cellular RNA was extracted from tumor cell samples, using the well known guanidine-isothiocyanate/cesium chloride methodology, (see, e.g., Davis et al., Basic Methods in Molecular Elsevier, NY (1986), pp. 130-135, which is not repeated here. Next, cDNA was synthesized, using 2 ug total RNA from the samples. Synthesis was carried out by extension with oligo dT(15), in a 20 μl reaction volume, in accordance with DeSmet et al, Immunogenetics 39: 121-129 (1994), incorporated by reference. After incubation for one hour at 42°C, the cDNA reaction mixture was diluted with water to 100 μl. PCR was then carried out using SEQ ID NOS: 5 and 6. Each PCR reaction was carried out, using 5 μl of cDNA (which corresponds to 100 ng of RNA), supplemented with 5 μl of 10xPCR buffer, and 1 μl of each variety of dNTP (10 mM), 0.5 μl each of 80μ M solutions of primers, 1.25 units of AmpliTaq DNA polymerase and water, to a total volume of 50 μl. This mixture was then heated to 94°C for five minutes, followed by amplification in a thermal cycler for 30 cycles (one minute at 94°C, two minutes at 63°C, two minutes at 72°C). Cycling was then concluded with a final extension step (15 minutes, 72°C). A 10 μl sample of each reaction was run on 1% agarose gel, and visualized using ethidium bromide fluorescence.

RNA integrity was verified, and samples containing strongly degraded RNA excluded, by carrying out a 20 cycle PCR assay, using primers specific for ß-actin, in accordance with Weynants et al, Int. J. Cancer 56: 826-829 (1994) incorporated by reference.

The results for tumors follow. The first column is the number of tumor samples tested, the second is the number which were positive for MAGE-B2:

Testicular seminoma	6	5
Non-small cell lung carcinoma	20	6
Melanoma	26	5
Breast	10	2
Sarcoma	10	1
Leukemia	10	1

With the exception of the positive leukemia, any tumor sample which was positive for MAGE-B2 was also positive for at least one MAGE-Xq.

Expression of MAGE-B2 was found in fetal and adult testis, but was not found in any normal kidney, liver, adrenal gland, skin, breast, brain, heart, ovary, prostate, cerebellum, peripheral blood lymphocyte, colon, stomach, lung, bladder, bone marrow or endometrium cells.

EXAMPLE 3

Additional experiments were carried out on cosmids D5 and 4965, which are discussed in example 1, supra. Specifically cDNA as disclosed by Muscatelli, et al, Proc. Natl. Acad. Sci USA 92:4987-4991 (1995), was subjected to PCR amplification. In these amplifications, the primers:

5'- GTGGTGTCCA GCAGTGTCT C -3' (SEQ ID NO:11)

and

SEQ ID NO:12

were used, to generate a 0.45kb probe. A second probe was then prepared using:

5' - AAT GTG TTG GGA GCC TAT GAT -3' (SEQ ID NO:13)

and

5' - ATT ATG TTG TGT GAG GTT CTT TCA -3' (SEQ ID NO:14)

to generate a 726 base pair probe.

The first probe contained 41 bp of exon 1, 105 bp of exon 2, and 300 bp of exon 3 of MAGE-B1, while the second probe consisted of the 726 bp at the 3'-end of exon 4.

Southern blotting was then carried out on both cosmids, using standard methods as can be found in, e.g., Lurquin, et

al, Cell 58:293-303 (1989). Any fragments of the cosmids which hybridized with the probes were cloned into commercially available vectors (ethyl pTZ18R or pTZ19R), and then sequenced.

The results of this work identified three sequences which showed significant identity to the last exon of MAGE-B1, as reported by Muscalelli, et al, supra. One sequence was identical to MAGE-B2, as described in Lurquin, et al, U.S. Patent No. 5,587,289, as MAGE-Xp2, and by Dabovic, et al, Mamm. Genome 6:571-580 (1995), as "DAM 6". This meant that there were two other homologous genes present in the cosmids.

EXAMPLE 4

In order to determine the precise positions, and complete sequences of the positives described supra, the portion of the Xp arm of the X chromosome, found in cosmids D5 and 4965, that includes the sequences of these hybridizing fragments was sequenced by "chromosome walking" (as described in Molecular Biology of the Cell, Alberts et al., Second Edition p.262-265).

A total of 40,352 kb was sequenced and this complete sequence is set out in SEQ ID NO:15. No further sequencing was carried out after this 40.352 kb sequence has been obtained because the start site and 5' UTR of MAGE-B2 was at the 5' end of this 40.352 kb sequence and the stop condon and poly-A signal of MAGE-B1 was located at the 3' end of this 40.352 kb sequence. At this point it was clear that all of the Xp hybridizing fragments from the Southern analysis (described in Example 3) were located within this 40,352 kb sequence obviating the need for any further sequencing.

When the sequence information obtained in example 3 was compared to the full, 40,352 bases of SEQ ID NO:16, the following was discovered:

GENE	POSITION IN SEQ ID NO:15
B2	3266 - 7979
B3	23546 - 25194
B4	29748 - 31474

B1 31403 - 39691

Within these sequences, further analyses showed that B2 contains two exons, at nucleotides 3266-3364, and 6278-7979, respectfully. The entire coding region is found at nucleotides 6283-7224, with a poly A signal being found at nucleotides 7961-7966.

As to the B3 gene, a single coding exon, at nucleotides 23546-25194 was found. The coding region consisted of nucleotides 23607-24647, with a poly-A signal at nucleotides 25152-25157.

The gene for B4 is thought to extend through to poly-A signal at 31822-31827, with the coding sequence being found at nucleotides 29808-30848.

The MAGE-B1 gene is the most complex of the four. The first exon, at nucleotides 31403-31474, is within the MAGE-B4 coding exon. Exons 2, 3 and 4 are found at nucleotides 33958-39691, i.e., at 33958-34062, 35057-35139, and 38088-39691, respectively. The coding sequence is found completely within the fourth exon, i.e., at nucleotides 38148-39191. The poly-A signal is at 39674-39679.

EXAMPLE 5

Comparison of the nucleotides in these sequences and other known tumor rejection antigen precursors, is set forth in Table 1, which follows. It can be seen that MAGE-B1, B2 and B4 form a closely related set, with about 80% identity while MAGE-B3 is about 70% identical with the others.

Further comparison reveals protein encoding regions corresponding to 347, 313, 346, and 346 amino acids for the MAGE-B proteins. These show anywhere from 49-68% identity.

Table 1. Sequence comparison of the human and mouse MAGE coding regions and proteins

% Nucleotide Identity

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	B1	B2	B3	B4	sB1	sB2	sB3
MAGE-A1	100	80	81	84	81	81	84	77	76	69	75	81	62	69	61	80	58	97	96
MAGE-A2	80	100	92	82	82	92	91	78	74	65	73	93	59	60	58	81	86	80	96
MAGE-A3	81	92	100	82	88	88	80	76	78	68	75	92	61	61	60	82	84	84	94
MAGE-A4	84	82	82	100	87	83	84	79	79	69	77	83	63	61	61	82	87	87	97
MAGE-A5	81	82	88	87	100	86	74	77	73	80	74	84	60	65	62	87	91	91	86
MAGE-A6	81	92	88	83	86	100	80	776	76	68	76	92	61	61	60	81	84	84	84
MAGE-A7	84	91	80	84	74	80	100	83	87	70	79	81	64	66	64	82	86	86	93
MAGE-A8	77	78	76	79	77	78	83	100	78	68	76	76	60	60	58	81	87	87	96
MAGE-A9	76	74	78	78	73	78	87	79	100	69	78	77	62	59	61	83	87	87	96
MAGE-A10	69	65	68	69	60	68	79	68	69	100	72	68	60	60	63	83	88	88	96
MAGE-A11	75	73	75	77	74	76	78	76	78	72	100	75	62	62	62	85	89	88	98
MAGE-A12	81	93	92	83	84	92	81	78	77	68	76	100	62	60	60	82	86	86	97
MAGE-B1	62	59	61	63	60	61	54	60	62	60	62	62	100	80	65	81	83	83	93
MAGE-B2	69	60	61	61	62	61	56	59	62	60	62	80	80	100	66	76	84	84	94
MAGE-B3	61	59	60	61	62	60	54	61	61	58	62	80	65	66	100	71	80	80	91
MAGE-B4	80	81	82	82	87	81	83	81	83	68	83	82	81	76	71	100	81	81	91
Smage-B1	69	65	64	67	61	64	68	67	67	66	66	66	63	64	60	81	100	99	98
Smage-B2	97	96	84	87	91	84	88	87	87	66	66	66	63	64	60	81	99	100	98
Smage-B3	96	96	94	97	96	94	95	96	96	66	66	67	63	64	61	81	98	98	100

% Amino Acid Identity

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	B1	B2	B3	B4	sB1	sB2	sB3
MAGE-A1	100	87	87	76	88	88	23	84	80	52	59	87	38	38	43	43	36	36	37
MAGE-A2	87	100	84	87	88	84	18	82	89	46	56	86	39	36	38	40	33	33	33
MAGE-A3	87	84	100	87	72	86	17	82	90	47	60	85	37	37	34	38	33	33	34
MAGE-A4	76	87	87	100	78	87	23	86	84	61	82	87	42	39	41	41	36	38	38
MAGE-A5	88	88	72	78	100	72	13	82	82	39	60	89	28	30	20	24	10	10	20
MAGE-A6	88	84	86	87	72	100	18	82	86	49	60	84	37	37	35	40	34	34	34
MAGE-A7	23	18	17	23	13	18	100	25	27	20	21	17	16	16	20	14	13	13	14
MAGE-A8	84	82	82	86	81	82	26	100	86	84	80	84	38	35	38	38	29	29	29
MAGE-A9	80	89	88	84	82	88	27	86	100	50	69	89	39	38	38	43	34	34	34
MAGE-A10	52	46	47	61	39	49	20	84	50	100	50	46	41	38	41	47	36	36	36
MAGE-A11	59	56	60	62	60	60	21	80	69	50	100	69	42	38	40	44	38	38	38
MAGE-A12	87	86	85	87	69	84	17	84	89	45	69	100	39	37	36	40	33	33	34
MAGE-B1	38	39	37	42	28	37	16	38	39	41	48	38	100	82	49	88	47	47	47
MAGE-B2	38	38	37	39	30	37	16	33	38	36	38	37	82	100	49	83	48	48	47
MAGE-B3	42	36	34	41	30	36	20	38	38	41	40	38	49	49	100	66	42	42	43
MAGE-B4	43	40	38	41	24	40	14	38	43	47	44	40	88	83	95	100	89	89	91
Smage-B1	38	39	33	38	19	34	13	28	34	36	38	33	47	48	48	62	100	100	97
Smage-B2	36	38	33	38	19	34	13	28	34	36	38	33	47	48	42	62	100	100	97
Smage-B3	37	33	34	38	20	34	14	28	34	36	39	34	47	47	43	61	97	97	100

EXAMPLE 6

In work reported by Muscatelli, et al, Proc. Natl. Acad, Sci. USA 92:4987-4991 (1995), MAGE-B1 from a cDNA library from testis was found to comprise two types, i.e.,

one included all four exons, and the other, exons 1, 2 and 4.

Experiments were carried out to verify this, using SEQ ID NOS: 11 and 12, set forth, supra, on a testis cDNA library, using RT-PCR. To carry this out, total cellular RNA was extracted, using the well known guanidine - isothiocyanate/cesium chloride method of, e.g., Davis, et al, Basic Methods In Molecular Biology, Elsevier Science Publishing Co., Inc., New York (1986). Samples (2μg), of total RNA were used for cDNA synthesis, via extension of oligo dt(15), in 20μl reaction volumes. See DeSmet, et al, Immunogenetics 39:121-129 (1996). The cDNA was incubated at 42°C, for 1 hour, and then diluted to 100μl with water. The primers set forth, supra, were then combined with 5μl of cDNA, together with 5μl of 10xDNA polymerase buffer, 1μl of each of 10mM dNTP, ad 1 unit of .DNA polymerase. Water was added to a total volume of 50μl. The mixture was heated to 94°C for 5 minutes, followed by amplification for 30 cycles (a cycle: 1 minute 94°C, 2 minutes at 63°C, and an extension of 2 minutes at 72°C). The cycling was concluded with a final extension step of 15 minutes at 72°C. Following this, a 10μl sample of the reaction was run on a 1.5% agarose gel, and visualized by ethidium bromide fluorescence. RNA integrity was verified, and samples with strongly degraded RNA, were excluded by carrying out a PCR assay of 20 cycles, using B-actin specific primers.

The results verified the previous findings, that there were two types of transcript which were present. The transcript containing 4 exons was far less abundant than the other.

The pattern of amplification products using SEQ ID NO:17 and 12 was also determined using RT-PCR on a testis cDNA library. In addition to a species comprising all 4 exons, a major species containing exons 3 and 4 was obtained.

Eighty-four tumor samples and tumor cell lines of various histological types were found to be negative for

MAGE-B1 expression when tested with primers whose sequences were located in the first and fourth exons (SEQ ID NOS:11 and 12). However, using primers whose sequences were located in the third and fourth exons (SEQ ID NOS:17 and 12), MAGE-B1 expression was detected in samples from NSCLC and mammary carcinoma and tumors of other histological type patients.

EXAMPLE 7

The pattern of distribution of expression of the MAGE-B genes was studied, via RT-PCR.

The protocol set forth in example 5, supra, was followed with some changes, as indicated herein.

Various combinations of primers were used, based upon the MAGE-B sequences. In addition to SEQ ID NO: 11 and 12, presented supra, the following primers were used for MAGE-B1:

5'-GAT CAT CCA GGA GTA CAA CTC GA -3' (SEQ ID NO:16)

5'-CCC GAG CGA GCT TAA GGA GT -3' (SEQ ID NO:17)

SEQ ID NOS: 11, 16 and 17 are sense primers corresponding to 1, 2 and 3, respectively, of MAGE-B1. One of these was used in combination with SEQ ID NO:12, in assays for expression of MAGE-B1.

For MAGE-B2, one of

5' -AGC GAG TGT AGG GGG TGC G -3' (SEQ ID NO:18) or SEQ ID NO:15, supra, together with SEQ ID NO:6, supra, were used. SEQ ID NOS:5 and 18 are sense primers for exons 1 and 2 of MAGE-B2, while SEQ ID NO:6 is an antisense primer for exon 2.

As indicated, RT-PCR was carried out, essentially as in Example 5, with the following exceptions. Forty cycles were carried out for MAGE-B1, while MAGE-B2 was assayed using thirty cycles. The cycle parameters given in example 5, supra, was modified as follows. When SEQ ID NOS:17 and 12, and SEQ ID NO:18 and 6, were used, a cycle was 1 minute at 94°C, and 2 minutes at 68°C, followed by the two minute extension. When SEQ ID NOS: 16 and 12 were used, the two minutes was carried out at 65°C.

The results are set forth in Table 2, which follows:

	Mage-B1 LUR171-1338 40 cycles	Mage-B2 LUR84-LUR85 and/or LUR162-LUR85 30 cycles
Surgical tumor samples		
Colorectal carcinoma	0/12	0/12
Gastric carcinoma	0/2	0/2
Leukemia	0/48	1/50
Myeloma	0/1	0/1
Melanoma	8/36	8/37
Skin carcinoma	1/4	0/4
Naevus (benign lesion)	0/8	0/6
Brain tumor	0/8	0/8
Neuroblastoma	0/2	0/2
Head and neck squamous cell carcinoma	0/12	2/12
Pleural mesothelioma	0/3	0/3
Small cell lung carcinoma	0/1	0/1
Non-small lung carcinoma	4/29	13/29
Sarcoma	1/11	2/11
Mammary sarcinoma	2/12	3/12
Prostate adenocarcinoma	0/6	0/6
Testicular tumor	8/9	8/9
Renal cell carcinoma	0/11	0/11
Bladder carcinoma	0/12	0/12
Cell lines		
colorectal carcinoma	0/5	0/5
Leukemia	0/3	0/3
EBV transformed B lymphocytes	0/1	0/1
Melanoma	2/9	3/9
Small cell lung carcinoma	0/2	1/2
Non small cell lung carcinoma	0/6	3/6
Sarcoma	0/2	0/2
Normal tissues		
Colon	0/1	0/1
Stomach	0/1	0/1
Liver	0/1	0/1
Bone marrow	0/1	0/1
Peripheral blood lymphocytes	0/1	0/1
Thymocytes	0/1	0/1
Skin	0/1	0/1
Brain	0/2	0/2
Cerebellum	0/1	0/1
Heart	0/1	0/1

Lung	0/1	0/1
Breast	0/2	0/2
Ovary	0/1	0/1
Uterus	0/2	0/2
Prostate	0/1	0/1
Testis	2/2	2/2
Adrenal gland	0/1	0/1
Kidney	0/1	0/1
Bladder	0/1	0/1
Fetal tissues: liver	0/1	0/1
brain	0/1	0/1
testis	1/1	1/1
placenta	0/1	1/1

Note that, in this table and the table which follows "LUR 171" is SEQ ID NO:17, "1338" is SEQ ID NO:12, "1339" is SEQ ID NO:11, "LUR 162" is SEQ ID NO:18, "LUR 84" is SEQ ID NO:5 and "LUR 85" is SEQ ID NO:6.

EXAMPLE 8

It is known that certain MAGE genes are inducible with 5-aza-2'- deoxycytidine, in both melanoma cells, and in different cell types which do not normally express the genes. See Weber, et al, Cancer Res 54:1766-1771 (1994); DeSmet, et al, Proc. Natl. Acad Sci. USA 93:7149-7153 (1996); DePlaen, et al, Genomics 40: (1997). Additional agents may also be used to induce MAGE genes.

In order to determine if the MAGE-1 genes parallel other genes in terms of inducibility, different types of cells were incubated for 72 hours in culture medium containing 1μm 5-aza-2'-deoxycytidine ("DAC" hereafter), in accordance with DeSmet, et al, supra. The table which follows sets forth the result.

	MAGE-B1				MAGE B-2	
	LUR171-1338 (exon3-exon4)		1339-1338 (exon1-exon4)		LUR162-LUR85 (exon1-exon2)	
	-	+DAC	-	+DAC	-	+DAC
Cell lines:						
MZ2-MEL	-	+	-	-	-	+
SK23-MEL	-	-	-	-	-	-

M1665/2-MEL	-	+	-	-	-	+
LE92.11-RCC	-	-	-	-	-	+
JAR	-	+	-	-	-	+
LB23-SAR	-	+	-	-	-	+
B-EBV	-	+	-	-	-	+
Normal tissues:						
PBL-PHA	-	+	-	-	-	+
Fibroblasts	-	-	-	-	-	+
Dentritic cells	-	+	-	-	-	+

"Nucleic acid molecule" as used herein refers to all species of DNA and RNA which possess the properties discussed supra. Genomic ("gDNA") and complementary DNA, or "cDNA" both code for particular proteins, and as the examples directed to isolation of MAGE coding sequences show, this disclosure teaches the artisan how to secure both of these.

The four MAGE-B genes are spread over 40,352kb in the 160kb X-linked critical region defined for the DSS (Dosage Sensitive Sex reversal) locus involved in sex determination (Bardoni et al. Nature Genetics 7:497-501 (1994)). This region is duplicated in patients with a male-to-female sex reversal phenotype. Genes in this region may be involved in X-linked disorders such as adrenal hypoplasis congenita and hypogonadism.

All isolated nucleic acid molecules which encode MAGE-B proteins, with the exception of MAGE-B1, are encompassed by this invention. This includes those nucleic acid molecules which hybridize to any of MAGE-B2, MAGE-B3, or MAGE-B4 under stringent conditions. As used herein, this refers to conditions such as hybridization with $5x10^6$ cpm/ml for 18 hours at 65°C, followed by 4, 20 minute washes at 65°C, with each wash using 2xSSC, 0.5% SDS and 1xDenhardt's solution, followed by two washes at 0.2xSSC, 1% SDS (20 minutes, each wash), and, finally, two washes at 68°C, 1% SDS, a varying concentration of SSC, each of these washes being for 20 minutes. The final concentration of SSC should be no greater than 0.5xSSC, more preferably it is 0.2xSSC, and most preferably it is 0.1xSSC.

Similarly, RNA molecules, such as mRNA can be secured. Again, with reference to the skilled artisan, once one has a coding sequence in hand, mRNA can be isolated or synthesized.

Complementary sequences which do not code for TRAPs, such as "antisense DNA" or mRNA are useful, e.g., in probing for the coding sequence as well as in methodologies for blocking its expression.

It will also be clear that one may manufacture biologically pure cultures of prokaryotic and eukaryotic cell lines which have been transformed or transfected with nucleic acid sequences which code for or express the MAGE-B molecules. Such cultures can be used as a source for tumor rejection antigens, e.g., or as therapeutics. This aspect of the invention is discussed infra.

Cells transfected with MAGE-B coding sequences may also be transfected with other coding sequences. Examples of other coding sequences include cytokine genes, such as interleukins (e.g., IL-2 or IL-4), or major histocompatibility complex (MHC) or human leukocyte antigen (HLA) molecules. Cytokine gene transfection is of value because expression of these is expected to enhance the therapeutic efficacy of the biologically pure culture of the cells in vivo. The art is well aware of therapies where interleukin transfectants have been administered to subjects for treating cancerous conditions. In a particularly preferred embodiment, cells are transfected with sequence coding for each of (i) MAGE-Xp molecule, (ii) an HLA/MHC molecule, and (iii) a cytokine.

Transfection with an MHC/HLA coding sequence is desirable because certain of the TRAs derived from MAGE-B may be preferentially or especially presented only by particular MHC/HLA molecules. Thus, where a recipient cell already expresses the MHC/HLA molecule associated with presentation of a TRA, additional transfection may not be necessary although further transformation could be used to cause overexpression of the antigen. On the other hand, it

may be desirable to transfect with a second sequence when the recipient cell does not normally express the relevant MHC/HLA molecule. It is to be understood, of course, that transfection with one additional sequence does not preclude further transfection with other sequences.

The term "biologically pure" as used in connection with the cell line described herein simply means that these are essentially free of other cells. Strictly speaking, a "cell line" by definition is "biologically pure", but the recitation will establish this fully.

Transfection of cells requires that an appropriate vector be used. Thus, the invention encompasses expression vectors where a coding sequence for the MAGE-Xp TRAP of interest is operably linked to a promoter. The promoter may be a strong promoter, such as those well known to the art, or a differential promoter, i.e., one which is operative only in specific cell types. The expression vectors may also contain all or a part of a viral or bacterial genome, such as vaccinia virus or BCG. Such vectors are especially useful in preparing vaccines.

The expression vectors may incorporate several coding sequences, as long as the MAGE-B sequence is contained therein. The cytokine and/or HLA genes discussed supra may be included in a single vector with the TRAP sequence. Where this is not desired, then an expression system may be provided, where two or more separate vectors are used where each coding sequence is operably linked to a promoter. Again, the promoter may be a strong or differential promoter. Co-transfection is a well known technique, and the artisan in this field is expected to have this technology available for utilization. The vectors may be constructed so that they code for the TRA molecule directly, rather than the MAGE-Xp TRAP. This eliminates the need for post-translational processing.

As the foregoing discussion makes clear, the sequences code for "tumor rejection antigen precursors" ("TRAPs") which, in turn, are processed into tumor rejection antigens

("TRAs"). Perhaps their most noteworthy aspect is as vaccines for treating various cancerous conditions. The evidence points to presentation of TRAs on tumor cells, followed by the development of an immune response and deletion of the cells. The evidence in the art shows that when various TRAs are administered to cells, a CTL response is mounted and presenting cells are deleted. This is behavior characteristic of vaccines, and hence TRAPs, which are processed into TRAs, and the TRAs themselves may be used, either alone or in pharmaceutically appropriate compositions, as vaccines. Similarly, presenting cells may be used in the same manner, either alone or as combined with ingredients or yield pharmaceutical compositions. Additional materials which may be used as vaccines include isolated cells which present the TRA molecule on their surface, as well as TRAP fragments, mutated viruses, especially etiolated forms, and transformed bacteria. "Fragments" as used herein refers to peptides which are smaller than the TRA, but which possess the properties required of a vaccine, as discussed supra. Another vaccine comprises or consists of complexes of TRA and HLA molecule. Vaccines of the type described herein may be used preventively, i.e., via administration to a subject in an amount sufficient to prevent onset of a cancerous condition.

The generation of an immune response, be it T-cell or B-cell related, is characteristic of the effect of the presented tumor rejection antigen. With respect to the B-cell response, this involves, inter alia, the generation of antibodies to the TRA, i.e., which specifically bind thereto. In addition, the TRAP molecules are of sufficient size to render them immunogenic, and antibodies which specifically bind thereto are a part of this invention. These antibodies may be polyclonal or monoclonal, the latter being prepared by any of the well recognized methodologies for their preparation which need not be repeated here. For example, mAbs may be prepared using an animal model, e.g., a Balb/C mouse or in a test tube, using, e.g., EBV

transformants. In addition, antiserum may be isolated from a subject afflicted with a cancerous condition where certain cells present a TRA. Such antibodies may also be generated to epitope defined by the inter-action of TRA and HLA/MHC molecules.

Review of the foregoing disclosure will show that there are a number of facets to the system which may be referred to as "tumor rejection antigen presentation and recognition". Recognition of these phenomena has diagnostic consequences. For example, the existence of specific CTL clones, or antibodies to the TRA makes it possible to diagnose or monitor cancerous conditions (explained infra), by monitoring the CTLs in a sample from a subject, binding of antibodies to TRAs, or the activity of anti-TRA CTLs in connection with subject samples. Similarly, the expression of nucleic acid molecules for TRAPs can be monitored via amplification (e.g., "polymerase chain reaction"), anti-sense hybridization, probe technologies, and so forth. Various subject samples, including body fluids (blood, serum, and other exudates, e.g.), tissues and tumors may be so assayed.

A particular manner of diagnosis is to use an adaptation of the standard "tuberculin test" currently used for diagnosis of tuberculosis. This standard skin test administers a stable form of "purified protein derivative" or "PPD" as a diagnostic aid. In a parallel fashion, TRAs in accordance with this invention may be used in such a skin test as a diagnostic aid or monitoring method.

The term "cancerous condition" is used herein to embrace all physiological events that commence with the initiation of the cancer and result in final clinical manifestation. Tumors do not spring up "ab initio" as visible tumors; rather there are various events associated with the transformation of a normal cell to malignancy, followed by development of a growth of biomass, such as a tumor, metastasis, etc. In addition, remission may be conceived of as part of "a cancerous condition" as tumors

seldom spontaneously disappear. The diagnostic aspects of this invention include all events involved in carcinogenesis, from the first transformation to malignancy of a single cell, through tumor development and metastasis, as well as remission. All are embraced herein.

Where "subject" is used, the term embraces any species which can be afflicted with a cancerous condition. This includes humans and non-humans, such as domesticated animals, breeding stock, and so forth.

There are therapeutic aspects of this invention as well. The efficacy of administration of effective amounts of TRAPs and TRAs as vaccines have already been discussed supra. Similarly, one may develop the specific CTLs in vitro and then administer these to the subject. Antibodies may be administered, either polyclonal or monoclonal, which specifically bind to cells presenting the TRA of interest. These antibodies may be coupled to specific antitumor agents, including, but not being limited to, methotrexate radio-iodinated compounds, toxins such as ricin, other cytostatic or cytolytic drugs, and so forth. Thus, "targeted" antibody therapy is included herein, as is the application of deletion of the cancerous cells by the use of CTLs.

The terms and expression which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expression of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.

MAGE-B Cluster Sequence

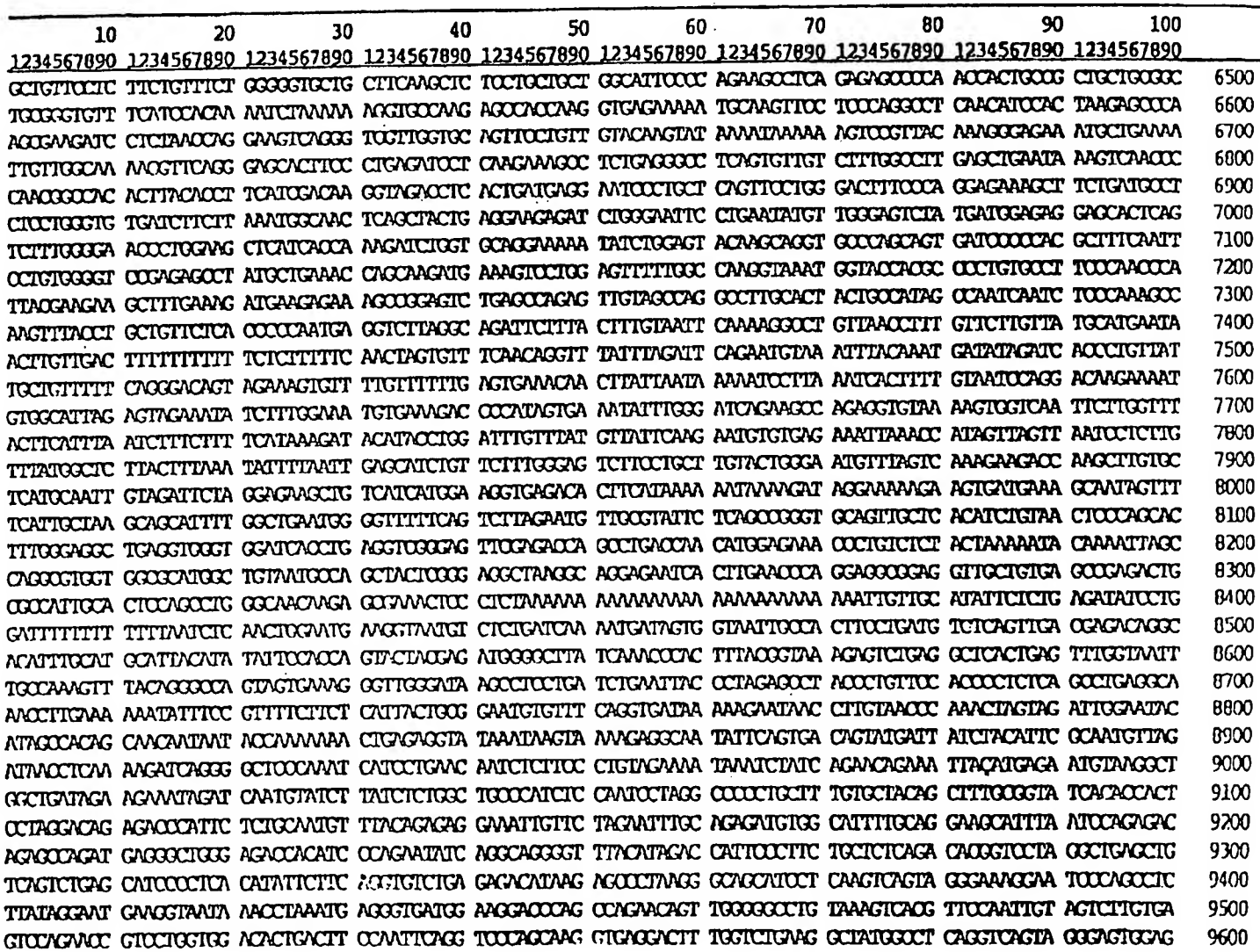
SEQ ID NO: 17

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
AAGCTTCAGA GTTGAGAGAC TAAGCCTATA ATGGAAAAGT GCTCTTAACT TTGGAGAAAA CTGGAAAGAA GCTCCACCAG GGGCAGCACT GGAAAGTGAC   100
TATGCTCTTG TCCCAGTGCA GATGAACACA GTCCAGAAAC TAGCTGCTTT ATATACATTG TATCCAGTGA GCTTCGGAGA AATAAGGATG ATACTCCCTT   200
AAGAAGAAAT GATAAGAAGC CGCTGGGTAC TCTGCACTGG GCTGGAAGAG TTAAGAGTCA ATTTCATTAA AAGGAGCATT CAAATTAGGC TATCATGTAC   300
ATAATTTGGC AAGCTCTAGG CAAGTTCTAG GTGTTTGGTG CTTGTTAAAT TAATGAAAAT AAAGGTGATT TAGATGGAAG TGACACTGGG AGGGAAGGAA   400
GTTGTTAGTC CTTGACGCAA ATGGGTTGTG TTGCAGCTAA ACTGGAGAAG TCTACCTCAC ATTGAATAAT TTATCATTTA ATGGGAAAAT ATTACTTGAT   500
TTTTCGTGCT AATCTGCATT TTGGCTTATT GGGTTTTCTT TCTTCTAGAA TGCTGCATAT TCTCTGACAT CTGTTGGATT AAGAAAAAGC AATCACAGTG   600
GGATGGGTGG TATCATCAAG TTAATCATAG GTAATAGCTG CCATGTTGAA ACCTCAATAT GGGCCAAGTA TTGTGCCAGG TGCTTTACAT ACATTAACCA   700
CATTTCACCA ATTCCACAAG GAAGGGGCTT ATCAAAGCTA TTTCACAGAT GCAGAGCGTG AGGCTCACAG AGTTTAGTAA TGTGCCTGAG TTCACATAGC   800
TGGTAAGTGG CAGGGCTGAA ACCAGACCCC CAGTCTGGAT TAATCTACAT CCCACACTGT TCCAGTCTTC CCATTCTGAG GCTAACCTCA TGTTGCTTCA   900
TTCATTTTTT CTCTCTATTT TTGAAAAATG TATCTCAGGC AATAAAAAGA AGGACTCTGT GCCCATACTA CTGGGTCAAA TATATTCCCA GACCAATATG  1000
AATGCCAAAA TAAATGAGAG GCATAAATAA AAAAAGAAA AATAGTACCT GGTGTCAATA CGTTCATCTA CACAAGCAAT ATCAGGTAAT CTCAAAAGAT  1100
CAGAAGCTCC CAAAATCATC CTGGACAGTC TCTTCCCTGT AGAAATAAA TGTGTCAGAA CAGAATTTTC GTGAGAATCT AAGTCTAGCA AGATGGAAAA  1200
AATCAGCATG TTTTTTTCTC TGACAGCCCA CTACCTGTCC TGGAGCTTCT GCCTTGTGCT TCAGCTTCCC TATCTCACAT CACTCCCGGA ACACGGCTCC  1300
TTCTCTGCAA GGTTTGCAGT AGAAAAACTG GTCATGTTTG CAGGGATGTG ACATTTCCCA GAAAGCCTTT AATGAAAGAT GATCCCCAGT TCAGAGGACA  1400
TTGGAAACCA ATATCTCAGA ATAATATGAGG TTATGCATAG ACTGACTCTC ATCTGCTCTC AGACCTAGAG AGCTCGAGAC AGAACTGTCA GGCTGAGCAT  1500
CCACTCATTT AATCTTCAGA CCTCTCATGG ACATGGCCTA AGAAGGTCAA CTTCATGTCA GTAGATGCAG TTCAGCCTCT GATAGATATG AACGTGAGGA  1600
CCCTGAATGA AGATGATGGA AGGACCCAGC CAGAACTGTG GGGTCCCAAA AAGTCCAGCT CAGGCTGTCA GCCCTTCAAA GTCCAGAACA GCCCTGGCTA  1700
AATGTACCTC ACCACAACTT CCAGTTCAGA GGTCGGCAGA GGGTGAGTAC CTTGATCTGA TGGGAACAGC CTCAGGTCAG TGGAGGATGG AGTCGCCCAA  1800
GACTGGCCAG GTGTCATGGT GAGGACCCTG AAGAGAAATA AGAGAACTGT TATAGGACCA ACGAGTTTCT GTGTCAGCTG CAAAGTAACA GACCCACTAT  1900
AATAAAACAG CAAAGTTTCC AGCAAAGAGT TTAATGATTG CAGGGTGCCA AGTGAGGAGA TGAGAAGAGA CTCTCAAATT CATCTCCCCA AGGAGTTCTG  2000
GGCTGGGATT TGCAAGGAGA TCATGGAGGG CGAGCGGCTG GAGAACTGAG GTTGTTGATT CGTTGGGGTA AAGGGGATGA AATCATCAGA ATGTGGAAAC  2100
TGCGTTTTTT TTTTTTTAGT CAGCTCCTGT GGGGTCTTTT GGACCAGCTG ATGTCAGTAG GGTCCTTTAG AGCAGCTGGC ATCAGTGGGG TCCTTCAGAC  2200
CAGCTGAGTC ATTAGAATCA TCAGTATGCA GGACCTTAAG GAATATCTCA AAGAAAAAAA GTTTCATTAT GTTAAAATTG TTATCTATAG AGCAATTAAG  2300
GGCAACTATA ATCTAGTAAC AGGATCTACA TGATTCTGGG ATAATAGGCA CCAAACTACT ATGAGGAAGC AGGTCAGAGA GCAGCTGACT TCATGATTAA  2400
TGCTGAATGT TCTGCAAGCT TGGCGTATCT TCATTTATCT CCCTTCCCTC TTCCCTGATT AATTTTATAC AGTTCATAGG GGCAGTTTCA AAACCACCCA  2500
CCCAGAACAG AGGAGCCCAC AAATTCCACC AATCTGTCAC CCTTGGGGAG CCCAGAAGAG AATTATGAGG TGGAGAAATC CCATCAATAT GCCTCGTGTG  2600
TTTCACAGAG GTGAAGACCT TGGACTAGTG GAAATGGCCC TAGGTCAAAG AGACAAAACC CACGCCCTAA CAGGAATCAA GGTGAGGATC CTAAGTGTTA  2700
CGAGGGGGCT CTCCCCCTTG CAACACCAGC AGAGGCGGCC AAACAGAGCC CTCGCCATTG TTAGCACCGA GAGCCCCAAA ATAAGTATCT CTGAATCCCT  2800
TCATTCCACC TGTATAGTCT CGGGGAAGGA AGGGCCTTTG TCTGAAGAGG GTGACCCAGT TGTGCAGAGA GAGAGTCTTG GCTTTCACGG GAATCAAGGT  2900
GAGGACTCTG AGGCGGGATG AGAAGACCTC TCCCCAAAAA AGGCACATTC ACAGAGCCCT GCCGCTGCTG TCAGGCCTGT GAGGCCAGGC AGGGGTGCCC  3000
TGTTTGGCAC GCTTAGATTT CCACAGTGGG GGCTGAGGGA GGTGGGGGTA TTGTTTGGAG GCTGCCGGAT TTGGGTCAGC ACGCATATTC GTCCCAGGCT  3100
GCTAGATACT GAGGTGAGGA CCCTAGTGGA GACGAAGGGA CCAGCAACGC TAGAACAGTG ACGTCCGGTA GCGTCCAGCC GTCAGCCCCT CAGACGCCAC  3200
```

MAGE-B Cluster Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
GGGCTGCCGG ATGTGAGTCA TCCTGACTTC CGCTTTGAAA AAAAAGACCC GAGCGGATGT GGCTCATCCT GACTTCCGCT TTGGAGGCGA GGACCCGAGC  3300
GAGTGTAGGG GGTGCGGCGT CTGGTCAGCC AGGGGTGAAT TCTCAGGACT GGTCGGCAGT CAAGGTGAGG ACCCTGAGTG TAAACTGAAG AGACCACCCC  3400
CACCCTGTAAC AAAGAGGGCC CCACTAAGTC CCGCTTCTGC ATTTGGTCCT GAGAGGCTCC GGTAAAGCCG TCCGGCAATG TTCCACCTGG AAAGTTCCAG  3500
GGCAGGGGAA GGGTGGGGGG AGGGGCAGTC GCGGGGGAAG GAGGTTTGGA CGCAGGGAAT AGGCCTCATT CTGCACGTAG GGTGGGTGTA GGCCCTAACT  3600
GAAATCAATT TGAGGCCCCT AAATGTGGAC TGAAGAGAAC ATCTCCTACC CTTAACAAAG GTGCCCTCAC TAAGTCTCGT CCCTGAAGAC GGCTCTGGGA  3700
GGCCCCAGCA AAGCTGTGCC TGGAAAGTCT CAGGGAGGGG AGGGCCTTGA ACCTAGGAAG CAGCCCTGCT TCTGCTCATA TGGCATCAAC CTAAAACCTA  3800
ACTGGGGTTA AGGTGAGGAC TCTGAGTATT AATGAGCGGA CCTCTACCCA AAAGAGGAGT CATGCTGGGC AGTTCCCAGG CCGAGGTGAC ACTGGGTTTT  3900
AGATGGTCCC AGGCTGACAT GACAAATGCA GAGCAACCTG ACTTCCTCTC CCCGGACTAA AGAAGTTGAG GGCTACATCC AAGACCCACA TCTGCTGCCA  4000
GCAATAGGAA GGTCAGGGCA AAGCTAGCAA GCTGAGATGC TCTCTAGTTT CCTCTGGAAG GTGGTCTCAG GGAGGTGAGG GTCTGGTGTA TGGGGTAGGT  4100
CTAATGCGAC AGAGAGGCCC AGTCTCTAAC AGGAAGCAAG GATGTGACAC TGAGTGATGA GGAGGGGACT CCAACCAGAA AGAAGTGCCA TATAGAGCCC  4200
ACTCTTGTTG TCAGACCAGG GAGACCTGAG CATGGCATGT TGAGTTGTAC TCACTCTCCC CAAGACATCT CAGAGAAGTG AACGACCCCA TTAAAGAGGG  4300
CAGCCTGGCC CGGCGCCGTG GCTCACGCCT GTAATCCCAG CACTTTGGGA GCCGAGACA GCCGGATCAC GAGGTCAGGA GATCGAGACC ATCCTGGCTA  4400
ACACGGTGAA ACCCCGTCTC TACTAAAAAT ACCAAAAAAA AAATTAGCCG GGAGCGGTGG CGGGCGCCTG TAATCCCAGC TACTCGGGAG GCTGAGGCAG  4500
GAGAATGGTG TGAACCTGGG AGGCGGAGTT TGCAGTGAGC CGAGATAGAG CCACTGCAGT CCGGCCTAGG CGAAAGGGCG AGACTCCGTC TCTTAAAAAA  4600
AAAAAAATAA AAAGCCAGCT TGAGGTCAGC AGAGGAGCGG TTTCCAGGTT GTGCCAGATG TTATGATAAG AACTCTGACG ACTGCAGGGG GCTCCCACCC  4700
CATATTAGTG GAGCCACACA TCCTCAGTAC TGTCAGCTCT GAGAGACCCC AGGCAGAAAT GTGAAACAGA GTGCCCATCA GTTCCCACTC AAGGGTAACA  4800
GCGAAATGAG AGTTTAATCT GAGGGTGTGA TCACTTGTCA ACAGAAGTAA TGTACTAATT GGTCCATTCT GGCATTACTA TAAATACCCA AGACTGGGTA  4900
ATTTACAAAG AAAAGAGGTT TAATTGGCTC ACAGTTCTGC AGGCTGTACA GGAAGCATGA TGCTGGCATT TGCTCAGCTT CTGGGAGCAG GGCTCAGGAA  5000
ATTTACAATC ATGGCAGAAG CGGAAGCGGA AGCAGACATG TTACATGGCC AGAGCAGGAT CAAGAGGCCA GCGGAGGCGC TACACACTTT TAAATGACCA  5100
GATCTCATGA GAAGTCATCG TTAAGGCGAA TCGTGCTAAA CCATTCATGA AAAACCTGCC CTCATGATCC AGTCATCTCC CACCAGGCTC CACCTCCAAC  5200
ATTCAGGATT ACAGTTTGAC ATGAGATTTG GTGAGGATAC AGATCCAAAC CATATCAGCG AAGAATCTTA GACCCTCTCA TGAGTCAAAT TTAGAGGCTG  5300
AGTTCGGACT GTATGTGGGA CCATCTATCA CAGCGCAGTT ATCTGCCTGC CACAGCTTTT ACTTTTGGGA AGACCACATT TGGTCAGATA AGGCCACCCT  5400
CACTTCCTCC TATCGGTTCT CAGGGATGTT CACTCTTACC ATGACGACAT AGGCCTCAGT TCAATAAAAA GGAGAGCTCC TCTACCAGGA GTCAAATTGA  5500
CGAATCTGAG AATTCAAGCA ACCACTGAAC CCTTAACAGT GGAGACATCA AAGAATCCAG TCCAACCCTT TATTTTCAGC CCTAGAAAGC TCAGGAAAGG  5600
GTTATCAGGT TTAGTGTCCC CCTCTCTTTT TTATATAAGA TCTAAGAGAG GTGAGATTCT TAGTCTGACG GTGTAGCCAC CACTCAGCAG AAGGGGCTG  5700
GTGCCAAGCC CTACATGAAG CCAAGGTCGG GACCTTGAAT GAGAGCTGAA GACTCTAATG ATTCCAAGAC ATCAAAGACC CCACTGAACT CTCAGCACTG  5800
GGTGTCACTT CTCAGGGTGG TGGTTGAGGC ACCCCTTCAA TTTCTTCTCA TGAGTCCCAG GGTGCTTTGA GGTGTCAAAT TCACAGTAGC GGAAGGAGT  5900
GGCTCAGGCC TGCAAGACTT CATATGATTA TCCTAAAAGT TAACTGATAA AACCCCAAAC ACCAGAAGAC AGGAGGCGGC ACTGCCAGTA CCTCCTGTCA  6000
CCATAGTAAG CCCAAGGAAG GGCTGACAGA ATGCAGTCTG AAACTCACTG TAGGGGTTCT GTGGTTCTTC TCTGGTGTCT CAGGGGAAAG GATAACCAGG  6100
ACGACAGGAG CCTTGTGGCT TTCTAGAACA GTGCCTTCAG GGAAACTTGC AAAGGCAGGC TTCCTTCAGT AAAGCTAAGA TGGGGTCTTC CAGCTGAAGG  6200
TGCTCACAGA TCTCATTCTC CCATCTCCAG GTATACTAAC CATCTATTCT TCTGCCCACA TTTCTTGGTT TACCCAGCCA TCATGCCTCG TGGTCAGAAG  6300
AGTAAGCTCC GTGCCCGTGA GAAACGCCGC AAGCCCCGAG ATGAGACCCG GGGTCTCAAT GTTCCTCAGG TCACTGAAGC AGAGGAAGAA GAGGCCCCCT  6400
```

MAGE-B Cluster Sequence



10	20	30	40	50	60	70	80	90	100	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
GCTGTTCCTC	TTCTGTTTCT	GGGGGTGCTG	CTTCAAGCTC	TCCTGCTGCT	GGCATTCCCC	AGAAGCCTCA	GAGAGCCCCA	ACCACTGCCG	CTGCTGCGGC	6500
TGCGGGTGTT	TCATCCACAA	AATCTAAAAA	AGGTGCCAAG	AGCCACCAAG	GTGAGAAAAA	TGCAAGTTCC	TCCCAGGCCT	CAACATCCAC	TAAGAGCCCA	6600
AGCGAAGATC	CTCTAACCAG	GAAGTCAGGG	TCGTTGGTGC	AGTTCCTGTT	GTACAAGTAT	AAAATAAAAA	AGTCCGTTAC	AAAGGGAGAA	ATGCTGAAAA	6700
TTGTTGGCAA	AACGTTCAGG	GAGCACTTCC	CTGAGATCCT	CAAGAAAGCC	TCTGAGGGCC	TCAGTGTTGT	CTTTGGCCTT	GAGCTGAATA	AAGTCAACCC	6800
CAACGGCCAC	ACTTACACCT	TCATCGACAA	GGTAGACCTC	ACTGATGAGG	AATCCCTGCT	CAGTTCCTGG	GACTTTCCCA	GGAGAAAGCT	TCTGATGCCT	6900
CTCCTGGGTG	TGATCTTCTT	AAATGGCAAC	TCAGCTACTG	AGGAAGAGAT	CTGGGAATTC	CTGAATATGT	TGGGAGTCTA	TGATGGAGAG	GAGCACTCAG	7000
TCTTTGGGGA	ACCCTGGAAG	CTCATCACCA	AAGATCTGGT	GCAGGAAAAA	TATCTGGAGT	ACAAGCAGGT	GCCCAGCAGT	GATCCCCCAC	GCTTTCAATT	7100
CCTGTGGGGT	CCGAGAGCCT	ATGCTGAAAC	CAGCAAGATG	AAAGTCCTGG	AGTTTTTGGC	CAAGGTAAAT	GGTACCACGC	CCCTGTGCCT	TCCCAACCCA	7200
TTACGAAGAA	GCTTTGAAAG	ATGAAGAGAA	AGCCGGAGTC	TGAGCCAGAG	TTGTAGCCAG	GCCTTGCACT	ACTGCCATAG	CCAATCAATC	TCCCAAAGCC	7300
AAGTTTACCT	GCTGTTCTCA	CCCCCAATGA	GGTCTTAGGC	AGATTCTTTA	CTTTGTAATT	CAAAAGGCCT	GTTAACCTTT	GTTCTTGTTA	TGCATGAATA	7400
ACTTGTTGAC	TTTTTTTTTT	TCTCTTTTTC	AACTAGTGTT	TCAACAGGTT	TATTTAGATT	CAGAATGTAA	ATTTACAAAT	GATATAGATC	ACCCTGTTAT	7500
TGCTGTTTTT	CAGGGACAGT	AGAAAGTGTT	TTGTTTTTTG	AGTGAAACAA	CTTATTAATA	AAAATCCTTA	AATCACTTTT	GTAATCCAGG	ACAAGAAAAT	7600
GTGGCATTAG	AGTAGAAATA	TCTTTGGAAA	TGTGAAAGAC	CCCATAGTGA	AATATTTGGG	ATCAGAAGCC	AGAGGTGTAA	AAGTGGTCAA	TTCTTGGTTT	7700
ACTTCATTTA	ATCTTTCTTT	TCATAAAGAT	ACATACCTGG	ATTTGTTTAT	GTTATTCAAG	AATGTGTGAG	AAATTAAACC	ATAGTTAGTT	AATCCTCTTG	7800
TTTATGGCTC	TTACTTTAAA	TATTTTAATT	GAGCATCTGT	TCTTTGGGAG	TCTTCCTGCT	TGTACTGGGA	ATGTTTAGTC	AAAGAAGACC	AAGCTTGTGC	7900
TCATGCAATT	GTAGATTCTA	GGAGAAGCTG	TCATCATGGA	AGGTGAGACA	CTTCATAAAA	AATAAAAGAT	AGGAAAAAGA	AGTGATGAAA	GCAATAGTTT	8000
TCATTGCTAA	GCAGCATTTT	GGCTGAATGG	GGTTTTTCAG	TCTTAGAATG	TTGCGTATTC	TCAGCCCGGT	GCAGTTGCTC	ACATCTGTAA	CTCCCAGCAC	8100
TTTGGAGGC	TGAGGTGGGT	GGATCACCTG	AGGTCGGGAG	TTCGAGACCA	GCCTGACCAA	CATGGAGAAA	CCCTGTCTCT	ACTAAAAATA	CAAAATTAGC	8200
CAGGGTGGT	GGCGCATGGC	TGTAATGCCA	GCTACTCGGG	AGGCTAAGGC	AGGAGAATCA	CTTGAACCCA	GGAGGCGGAG	GTTGCTGTGA	GCCGAGACTG	8300
CGCCATTGCA	CTCCAGCCTG	GGCAACAAGA	GCGAAACTCC	CTCTAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAATTGTTGC	ATATTCTCTG	AGATATCCTG	8400
GATTTTTTTT	TTTTAATCTC	AACTCGAATG	AAGGTAATGT	CTCTGATCAA	AATGATAGTG	GTAATTGCCA	CTTCCTGATG	TCTCAGTTGA	CGAGACAGGC	8500
ACATTTGCAT	GCATTACATA	TAATCCCACA	GTACCTAGAG	ATGGGGCTTA	TCAAACCCAC	TTTACCGTAA	AGAGTCTGAG	GCTCACTGAG	TTTGGTAATT	8600
TGCCAAAGTT	TACAGGGCCA	GTAGTGAAAG	GGTTGGGATA	AGCCTCCTGA	TCTGAATTAC	CCTAGAGCCT	ACCCTGTTCC	ACCCCTCTCA	GCCTGAGGCA	8700
AACCTTGAAA	AAATATTTCC	GTTTTCTTCT	CATTTACTGG	GAATGTGTTT	CAGGTGATAA	AAAGAATAAC	CTTGTAACCC	AAACTAGTAG	ATTGGAATAC	8800
ATAGCCACAG	CAACAATAAT	ACCAAAAAAA	CTGAGAGGTA	TAAATAAGTA	AAAGAGGCAA	TATTCAGTGA	CAGTATGATT	ATCTTACATTC	GCAATGTTAG	8900
ATTACCTCAA	AAGATCAGGG	GCTCCCAAAT	CATCCTGAAC	AATCTCTTCC	CTGTAGAAAA	TAAAATCTAT	AGAACAGAAA	TTACATGAGA	ATGTAAGGCT	9000
GGCTGATAGA	AGAAATAGAT	CAATGTATCT	TATCTCTGGC	TGCCCATCTC	CAATCCTAGG	CCCCCTGCTT	TGTGCTACAG	CTTTGCGGTA	TCACACCACT	9100
CCTAGGACAG	AGACCCATTC	TCTGCAATGT	TTACAGAGAG	GAAATTGTTC	TAGAATTTGC	AGAGATGTGG	CATTTTGCAG	GAAGCATTTA	ATCCAGAGAC	9200
AGAGCCAGAT	GAGGGCTGGG	AGACCACATC	CCAGAATATC	AGGCAGGGGT	TTACATAGAC	CATTCCCTTC	TGCTCTCAGA	CAGGGTCCTA	GGCTGAGCTG	9300
TCAGTCTGAG	CATCCCCTCA	CATATTCTTC	AGGTGTCTGA	GAGACATAAG	AGCCCTAAGG	GCAGCATCCT	CAAGTCAGTA	GGGAAAGGAA	TCCCAGCCTC	9400
TTATAGGAAT	GAAGGTAATA	AACCTAAATG	AGGGTGATGG	AAGGACCCAG	CCAGAACAGT	TGGGGGCCTG	TAAAGTCAGG	TTCCAATTGT	AGTCTTGTGA	9500
GTCCAGAACC	GTCCTGGTGG	ACACTGACTT	CCAATTCAGG	TCCCAGCAAG	GTGAGGACTT	TGGTCTGAAG	GCTATGGCCT	CAGGTCAGTA	GGGAGTGGAG	9600

MAGE-B Cluster Sequence

```
        10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
TCCCAGGAAG GGGGCAGAGT CAAGGTGAAG GCCTGGCATG TGGATATGGG GACCACTCAA CCCATAACTG GCTGGCGTGC CTCTAGTCAT TCCTTTGCTA  9700
TTAGCCCTAG GAGGCCTAGG GAGGAGCTGG CAGGCCCAGG TGCCCGTGTC TTCTTCCTGG AGAATCTCAA GGAGATGAGA GACTTGATTT ACAGGCTGGC  9800
CTCCCAGTTA AGCAAAGAGT GGAATCCCAG ACCATGATAG GTATTAAGGT AAAAATCTCA TATGAGGAAA GAGGCAAACA ACTGCTAGAG AGGTGAACAG  9900
AACAAGGACG TCCTTAGAAA GTTATGCCGG GCCAGACGCG GTGGCTCATG CCTGTAATCC AGCACTTTGG CAGGCCGAGG TGGGCGGATC ATTTGAGGTC 10000
AGGAGTTCAA GATCAGCCTG GCCAGCATGG CAAAACCCCG TCTCTACTTA AAAAATACAA AAATTAGCTG GGCGTGATGG TGGGCGCCTG TAATCCCAGC 10100
TACTTGGGAG GCTGAGGCAG GAGAATTGC TTGAACCTGG GAGGCGAAGC TTGCAGTGAG CTGAGATCAC GCCACTGCAC TCCAGCCTGG GCAACAGAGC 10200
GAGACTTTAT CTTAAAAAAA AAAAAAGAAA AGAAAAGAAA GTTATGCCCT GGGAGGCCAT AAGCATAGCT TTCAGGCTGA GGTTACCCCA CATTGGCTGA 10300
GGGGAGGGGA GTGGTCTCAG GGACAGGAAG TCCTTGGACT AATAGGAGAG GCCCCATTCA GTAGAGGGAG GAGATTGGCT CCTTAACAGG ATTCAACATG 10400
ATGACACTGA GTGACAATGA GGGGACCCCT CTCAGAAAGA AGAGGGTCAC ATAGAACCTT GTCCCTATTG TCAGACCTGA GACCTGGACA TGATGACATG 10500
ATGTATCATA CTCACTTCTT CCCAGGTAAT CTGGCGGAGG TGAAAAATAT TAACAAAGTG GGCAGCCTCA GGTCAGTAGA GAGAAGATTT CTAGGTTGTT 10600
GCAGAAGTAA ATGTGAGGAC CCTGAAGACT GAGGTGAACA CCTACCCCAT AACAGTGAGG GTAACACAGA TTCCTTCCAC TACTGTTAGC CTAGGGAGAA 10700
CATGAGCAGT TATGGCAATA TGAAGCAAAT TTTACTTCTT CTTAGGGTGT ACCAGGAAAT TGGTCTTCAA TTCCCTTCAT ATGGGGAGAC AGGCCTCAAG 10800
TCAAGACAGG GAGGATTTGC ACAGGGTGTG ACAGGAATAA GGATGAGGAA ACTGAAGAGT ATGGGACCAC AGACCCCACA TCAGTAGGAG CCACACAGAA 10900
TCCTCCCTAC TGTCAGCCCT CAGAGACCCC AAGCAGAAAT GTCAGGTGGA GTTTCCCATC AGGTCCTACT CAAGGGTAAC AAGGGAAGTG AGGGTCTTGA 11000
TCTTAGGGTG GTAGTCTCAC ATCAACAGAA AAAGAAATCT TAGACCTGTC CAGAAGTTAA AAATTTAGGA CCCTGAGTTG GGACCCTGCA TGGAACTACC 11100
CATGATAGGG CCTTGAACTG CCTGCCACAG GTTTCATTCT TGTGAGACCA TTGGCAGGTA TATAGCCAGA TGAGGCCATC CTTGCTTCCT CTTGCTGGAT 11200
CTTAAGAAGG TGTGGGCCTC ACTTTGAGGA GATGTCATCA GGTCAAGAAG GAGGGGATCC CCAGGCCCTT CCAGGAGTAA AATCAGGGAA TCTGAGTGGA 11300
GACTGATGGT ACAACATAGC CTTGAATAGA GGGGACAACA AAGTGTCATG CCCTACCCCT TGATATCAGC CTTAGAAGAC TCAGAGCAGG GCTGTCAGGG 11400
GAGGCACCTC ATCACTTCTT TATATACGGT CTAGGGGAGG GGAAATCTTT GGTCTGAAAG TTCAGCTACC AGTCAGCAGA AGAGATCCAA TACATGCCAG 11500
TCACTGTGCC AAGTGCTTTC CATACATTAC CCACATCCTG CCAGTTCTGC CAGACAGGGC TTACCAAACC CATTCCATAG ATGCAGAACC TAAGGCTCAT 11600
AGGATTTAGT AATGTGCCCA AGTCACACTG CTAGTAATGA CAGGGCTGGG ATGAGAACAC TGGTCTGAGT TAGTCAAAAG CCCACACTGT GCCACTCCTC 11700
TCAGACAAAG GATGACCTTG TTCCTTAATT AACTTTTCCT CTCGTTACTG CAAAATGTAC CAGAGGCAAT AAAAAGAAAG ACTGTGGCCA AACGATGGAC 11800
TGTAATATAT AGCCAGAACA ATAAAAATAT CAAAATAAAT GGAAGGTTTG AGTAAGGAAA AAAATAATAT TTTGATGACA TAATTTACAT ACATATGTGA 11900
TGTCCAATAA ATTCAAAAAA TGAGGAGGTA CAAAAAAATC ATCTGAGCCA CTTCCTTAGA AAAAGTAAAT ATATCACAAC AGAAGTTACA TGAGACCCTA 12000
AATCTGGCTG GGCAAAAATA GGAGATGGAC ATTTGCTTTT TCTCTTATAG CCTACCACCT ATCCTGGCCC CCTGCCTTGT GCTGTAGATT TCCCATATTG 12100
CATCACTCCT AGGGCAGGGG CCTGTTCTCT GAAAGGTTTG CAGAGAGGAC ACTGCTCAAG AGTTTCGCAG AGATGTGGCA TTTCCCAGGA AGCCTTTAAT 12200
CCCACTGCAA GAGCCTGTAT GAGGACATAC ATAAATCAGG ACTGGAAAAA CAAACACTGC AGAATATTGG TTATTACCAA AAGACTATCC CCTGCCTGCT 12300
TTCAGACCTA GAGGGTCCCA GGCAGAACTG TCAGGATGAG CATCCTCTCA TTTATTTTTC ACAGTTGTCA GGAACTTGAG AGCTTTAGCC TAAGGAGGGC 12400
AACCTTAGGT CAGGAGAGAG ATAAGACCCA AGATCTGTTA GATACAAATG CCAGATCCCT TAAAAGTAAT TTAGAAAGCA CCCATCCAGA ACACCGGGGT 12500
CCCGAAGAGT TCAGGCCCTG CTGTTGACCC TTGGAGGTCC TGAATTCAAT GGTTGGATGA AATATAGAGA GGCCTTGAAA GATGAAGCAG AGAGAGCCTG 12600
AGCCAGAGCT GCAGCCAGGG CTGTGCCTACT GCCATGGCCA TGGCACATAC CAGAGCCACG TCCAGCTGCT CCTCCTACAT CTAGTGAGAT CTGAGACAGA 12700
TTCTTCACTT TGTAGTTGAA AAGATAAGTC AACATTCTAA GTAGTGGAGA GTCAATTTTG ACCTAGGGCA AACATATTGT ATGTCTTATT TTTGTTTTTG 12800
```

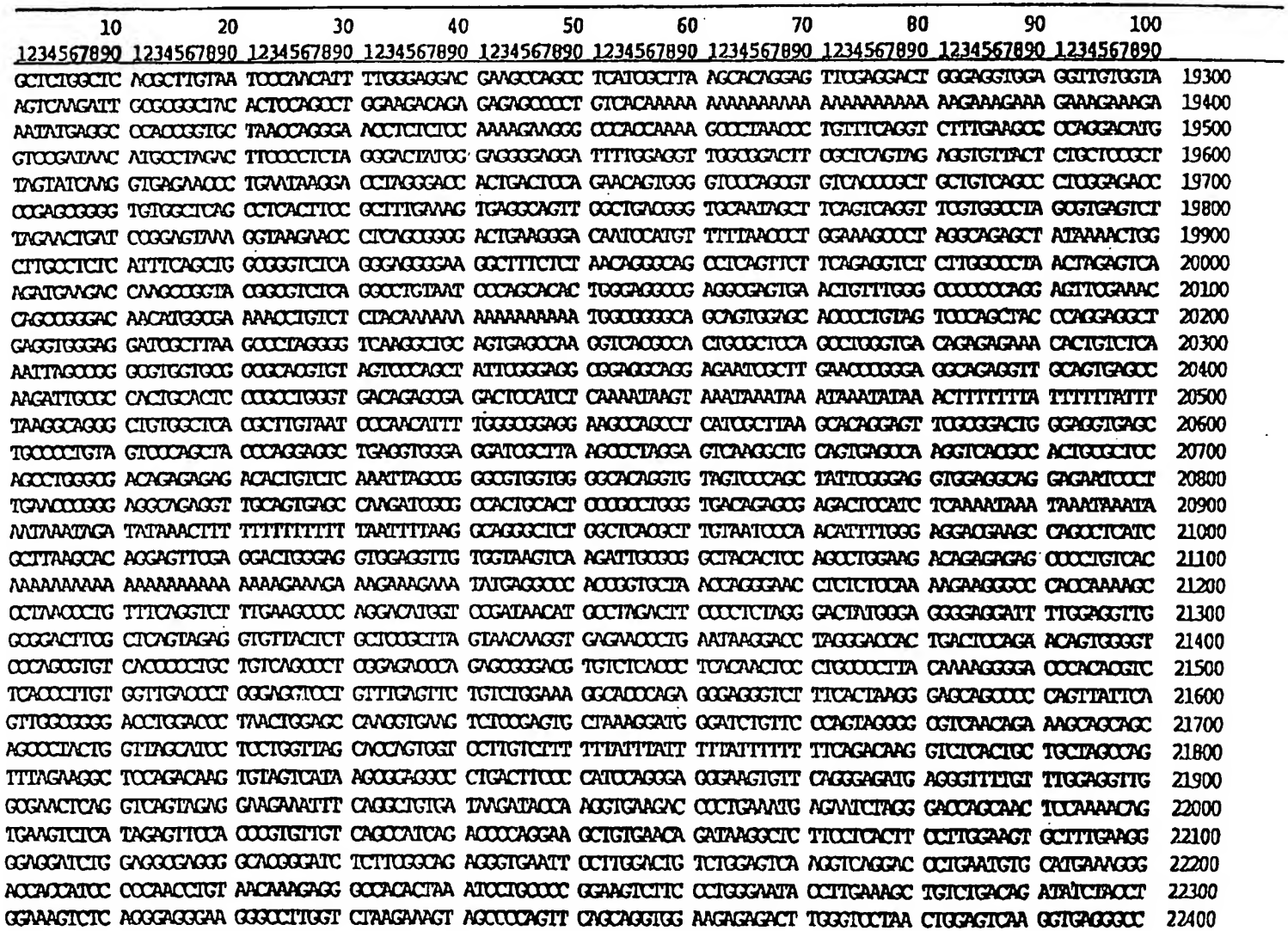
MAGE-B Cluster Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
CTCTACTTGA ATAATTGGNA GATGTATCTT TTTTATTTTT GGTACTTTTA AAATGTATTC ATTTTAATAG ANGATTTATT TAGCTTCAGT ATCTGTGTTT 12900
ATGNATANCA TGGATAACAT ATTTATTTCT GTTTTCCATA TATAAATGTA AAAGTGCTGG TATTTTTTTA TCANCAAACT GAAAATCCTT AGGTCTCTCT 13000
TTGTGGTCCA GANCAAGATA ACATAGCATA TGNATNAGGA CTTTTTTATT GTATNAATTT AAGTTGTACA GTAGGACGTT TTGATATACA TATATGCCCA 13100
TGCACCNCAT GATGATGTCT CAATCAGGAC CACATATACA ATGGCAGTCT CGTAAGATTA TAATGGAGCT AAAAAACTGT ATTGTTTACA TGTGTGTGGT 13200
GATGCTGGTG AAAACAAACC TACTGCACTA TCAGGCCTAT AAAAGTATAG CACATACAAT TATGTGCAGT ACATAGCGCT TGATACTGAT AATNAACAAC 13300
AATGTTACTG GTTTATATCT TTATTATACT ATAGTTTTTT TTAAATCAGG TTGGTGCAAA AGTNATTACA TTTTTTGCCA TTAAAAGTAA TAAGAATTGA 13400
GATGGGGTAG TAGTAGTAAT AATATTCTAG TAATACTAAA TAAGTAGTAA TATTANTAGG TAGTACTATT AAAATATTAA TATTAATTAA TAAGTAGTAA 13500
TATTANAAGT AGTAATAATT GGGACGGAGT CTTCCTCTGT TGTCCAGGCT GGGGTGCAGT GGTGCAATCT CAGCTCATTG CAACCTCTGC CTCCTGGATT 13600
CAAGCATTCT CCTACCTCAG CCACCTGAGT AGCTGGGATT ACAGGGGCCT GCCACCACGA CCAGATAATT TTTGGACTTT TAGTAGAGAT GGAATTTCAC 13700
CATGTTGGTC AAGCTGGTCT TGAACTCCTG ACCTCAAGTT ATCCACCCGC CTCATCCTCC CAAAGTGCTG GGATTACAGG TGGTACTTTT ATTATTATTT 13800
TAGAGTGCAC TCCTTCTACT AAAAAAAAAA ATGTTAACTG TAAAAACAGT CTCAGGCAGG TCCTTCAGGT TGTATTCCAG AAGAAGGTGG TGTTATCATA 13900
AGAGATGACA GCTCCATGCG TATTATTGTC CCTGAAGACC TTCCAGTGGG ACAAGATGTG GAGGTAGAAG ACAGTGATAT TGATGATCCT GACCTTGTGT 14000
AGGCCTAGGC TAATGTGGGT GCTTTTGTCT TCATTTTTAA CAAAAACGTT TAAAAATTTA AAAAGTAAAA ATAGAAAAAA GCTTATAAAA TAAGAATATA 14100
AAGAAAGAAA ATATTTTTGT ACAGCTGTAC AATGTGCTTA TGTTTTAAGC TAAGTGTTAC TACAAAAGAG TCAAAAAGCT AAAAAAAATT AAGAAGTTTA 14200
TAAAGCAAAA AAAGTTACAG TAAGCTAAGG TTAATTTATT GTTAAAGAAA GAAAACTATG CTTGAGAATT TAGTGTAGTC TAACTGTACA ATGTTTATTA 14300
AGTCTATAAT AGTGTACAGT AATGTCCCAG GCCCTCACAT TCACTCACTA TTCAATCACT GGCTCACCCA AAGCAACTTC CAGTTCTGCA AGCTACATTA 14400
ATGGTAAATA CTCCACACAG GTGTATGATT TTTAACAATC TTTTGTACCA TATTTTTTAC TGTACCTTTT CTGTGTTTAG GTACACAAAT ACTTAGCACC 14500
GTGTTACATT TGCCTATAGT TTTCAGTGCA GTCACATGCT GCACAGGTTT GTGGCCTAGG AGCAATAGGC TACACCAAAT AGCCTACGTG TGTAGTAGGC 14600
TATATACCAC CTAGGTTTAT GTAAGTATAC TCTATTATAT TTGCACAACA ATGAAATTGC CTAACAACGC GTTTCTCAGA AAGTATCTTC ATTGTTAAGC 14700
AATGCATTAC AGTACATGGT GAAATGCTTA CTACAGGTAA GCAATTTAAT ATATCCATTA TATCATATAG TTACCTTGTT TTGTGGTAAG AGCAGCTAAA 14800
ATCTACTCTT AGAAAATTTG CAGTATGCAA TACAATATTA CTAACTGTAG TCCTCATACT GTACTTTAGA TCTCTAGATT TATTCGTCTT ATAAAACTGC 14900
AACTTTGTAC TGTTTGACAT GCATCTCTGC ATCCCCTCCC CACCCTGCCT CTGGTAACTA CTGTTTTATT CTCTTTTTCT ATGTATTTAA CATTTTTTTC 15000
TTTTTCGATT TTACATATAA GGGATATTAA TAGCATGCAG CATTTTTCTT TCTGTGTCTG GTTTATTTCA TTCAGCATAA CGTCCCCCAA GTTCATCTGT 15100
GTTGTTGCAA TGGCAGAATC TCTTTCTTTT TCAAGGCTAA ATAACATTTT TTTAAATTTT ATCTTAGTTT CAGGGGTACA TGTGCAGGTT GTTATATAGG 15200
CAATTGCATG TCATGGGGGT CTGGGGTACA GATTATTTCA TCACCCAGGT AATAAGCATA GTACACAATA GGCAGTTTTT CCATCCTCAC CCTACTCCCA 15300
CCTTCCGCCC TTAAGTAGGC CCCAGTGTCT GTTGTTCCCT TCTTTGTGTC CATGTGTACT CANTGTTTAA CTCCCACAAA GAACATGCAG TATTTGGTTT 15400
TCTGTTTCAT TGTACATATA TACCACAATT CCTTTATCCA TTCATCCATT GGTAAACAAT TTTGTTGTTG CCGCATCTTG GTCATTGAGA ATAATACTGC 15500
AGTGAACATG GGGGCATAAA TATCTACAGG AGGTGATGAT TTCGTTTCCT TATGCCCAGA CAAGGGATTG CTGGGTCATA TGGTAGTCCA TTTTCAATTT 15600
TTTGAGAAAG CTCATACTGT ATTCTAGAAT GGCTGTACCA ATTTGCATTT CCACCAATAG TGTAGAAGCG TTCCATTTTC TCTACACTCT TGCCAACATT 15700
CATCTCCAGT GGGTTTTTGT TTGGTTGTTT TTGTTTTATT TTTTTATAAT AGCCTTCCTA ACTTGTGTGA GGTGATATCT CGCTGTGGTT TTGATTTGCA 15800
TTTCCCTGAT GGTTAGTGAT GTTGAGCACA TCTTTGTATA GCTATTGGAC ATTTTTATGT CTTCTTTGAG AAATGTCATA CAAGTAGTGA ATTTTCTTGA 15900
GTACTGATTA TGTTGGAAAT GTGATTGAAT ACTGTATTAA AAATTGTTGA GATCAAAAAA TTTTTTTAAA AAAGGTTGTC TATTCTTGGT TTGCCTAATT 16000
```

MAGE-B Cluster Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
CCTTTTAGTC TTTCTTCTTA TAAAATTAAG AGTTACATAT CTGGATTTGC TTAGATTATT AAAGAATGTA GGAGAGATTA AATCCTAATT TATTGGACTT 16100
CATACTCACT CTCTTGTTTA TTCCTTAATC ATTAATTGAG CATCAGCTCT TTGGAAGTCT TCATGTTAGT ACTGGGAATG TTTTCCCCAA AACAAATCAA 16200
TTTCTTGCCC ATTTAATTTT AGAGTCTGGG AGCAGCTGTC NTAAAAAAAA AAAAGATGAT GAGGTGCTCT CTAAAACTTA AAGAACAAGT ACAAAGGAGG 16300
AATGAGAAAA AGAGGAGGTT TGAGATGAGA GCAATCAAGT GTAAATGCCC TGAGGCAAGG CAGTTTGGAG TGTTAGGAAA CCTCAAGTCC CCCATTGGGA 16400
GGTAATTTTA AGGGAAACTA CATGGTGCGC TGGATGAGAC TGTGGGAGGG GCCAGGCCCT CAGATGGTGC CTCTCAGAGG TGTGAGACAA AGCCTGGAAT 16500
GGGAAGTAGT TCTTAACAGT TATTTTGTGC TCNCGGATAA ACCAGAGAGA NTCTGCNCCT GGGGCAGGAA TGGAATGTGC CCTGTGCTCT TGTCCCAGTG 16600
CAGCAGAACA TAGTCACAGT GCACAAACTA GGTGCTTTAT NCAGTTGGTG CATTCAGTGT TGAGAGNCTA AGCCCCGAAT GGGAAACTAT CCCTAACATT 16700
ATCACTTTCC TGTTGTTGGA AAACCAGAGA GAACCCCTAC CTAGGACAAA AGTGAAAAGT GTTCTATGTC CCTATCCTAG CACAGTCTAA TACAGTGCAC 16800
AACCTAGGTG TTCTATGTAC ATCATCTCTA GTGAGTTTCT GAGAAATAAG GGAGATGACA GCTTCAGGGG AGGTAAGATG CCCAGAAGCC ACCGTGCTGG 16900
CACTCTTTGT CCTGGGTTGG AGAATCAAGA GCCCGCTCTA TTAAAAGGAC TTTCAACAGG GGTGCAGCCA GGCATGGTGG TTCATGCCTG TAATCCCAGC 17000
ACTTTGGGAG GCGAGGCAGG TGGATCTCTT GAGATCAGGA GTTCAAGATT AGTCTGGCCA ACGTGGTGAA ACCCCGTCTC TACCAAAAAT·ACAAAAATTA 17100
GTTGGGCGTG GTGGTGGACA TCTGTAATCC CAGCTACTAG GGAGGCTGAG GCAGGAGAAT CACTTGAGCC CTGAGGCACA GGTTGCAATG AGAAGAGATA 17200
GTGCCTCTGC ACTCCAGCCT GGGTGACAGA GTGAGACTCC ATCTCAAAAA AAATAAATAA ATAAATAATA AAAATCAAAA CAGGAGTGGT TAACATAGAA 17300
GGGTGCCTAG GAGTTAGAAA AAAATATTGG TCATTGACAA AAATTTTGAG ACTTGAGTTG TATCCAATTG GAGAAGGCTC TTCCACAAAC ACTTTATAAA 17400
TTATATAATT TTCCTTGCTA AGCAGCATTT TGTTTGATTA TAATTTCTTT GTTTGGAGTT TGGTAATATG CTCAAGTTCA CAGGGTGAGA AGTGACAGGG 17500
CTGGGACTAG ACTCCTGTTG TGAATTCTTC TAGAGCCCAC ACTCTTCATT CTGTTCCTCT CTGCCTGGAG CAGACTTTGT GTTCTTTAAT TCACTTTTCT 17600
TCTCAATACT TCCAAATATA TCTCAGGGGT TAAAAAGAAG GATTCTGTGC CCAAAGAATT GGATTGGAAT ACATAGCCAA AGCAATAAAG AATATCAAAA 17700
TAAATGAAAG CTACAAATAA AGAGGAGAAA GAGATATTAT TCAGTGACAA TATGATCATC TACATATGCA CAAGTCAGAT AACCTCAAAA TCAGGGACCA 17800
CACAATCATC TTGAACAGCT CCTTCCCCTT AAGAAGTAAA TCAATTAAAA GAGAAGTCAC ATAAGAGGCT AGTGCTGGCT ATTAAAAGAG TAAATCAGCA 17900
TCATTTTTCT CTGACAACCC ACCAGCCTAC TTGGACTCCC TGCTTTGTGT TATAATTTCT TTACCTCACA TTAGCCCTGG GGCACAGACT TGCAAGGTTT 18000
GCAGTAGGAA CGCTGCCCAA GAGTTTGCAG GGATGTGGCA TTTCCCAGGA AGCCTTTATT CAGAGAGGCA AGAGAAAATA AGGAGGACAC CCATAAATAA 18100
ATACTGGAAA GACAACCACA CCAAAAGATA ACGGGGCAAT CTACCCAAAC CTCTTCTCAG CCTAGAATGC CCCAGGCAGT GCAGTCAGAC TGAGCATCCT 18200
CTCTTTTACC CTTCAGAGTT CTCAGTGAAG TGAGAGTTTT GACCTAAGGT GACCTCAGTT AACTAGAGGG AGGATCCAGG GTCTGGCAGA TATCAGGGAG 18300
ACCTTGAATG AAGAACAGGA GGAGCCAGAA CAGTGATGTC CATAGAAACC CGACTCCGCT GTCTGCCCTC ACAGGACCCA AACATCCCTG GCCTAATGTG 18400
GCTCATCAAA ACTTCAGCCT AGATCTCGGA GGACCTTGAT CTGAGCAGGA TTCATGGCAT GGAATGTGGA GTCCAGGACT GGCCAGCAGC GAAGTGAGGT 18500
TCTTGAGTGA GTAAAAAGGG AAAACTAAAC CCACAAATAA GGGGACTTCT CGGAGCCCAA TCCATATTTT TAACCCTGGA AAGCCCTAGG CAGAGCTATA 18600
AAACTGGCTT GCTCTCATTT CAGCCTGGCG ATCTCAGGGA GGGGAAGGCT TTGTCTAACA GGGCAGCCTC AGTTCTTCAG AGGTCTCTTG GCCCTAACTA 18700
GAGTCAAGAT GAAGACCAAG CCGGTGCCCA TCTCAGGCCT GTAATCCCAG CACACTGGGA GGCCGAGGCG AGTGAACTGT TTGGGCCCCC CAGGAGTTCG 18800
AAACCAGCCG GGACAACATG GCGAAAACCT GTCTCTACAA AAAAAAAAAA ATGGCGGGCA GCAGTGGAGC ACCCTGTAGT CCAGCTACCA GGAGGCTGAG 18900
GTGGGAGGAT CGCTTGAGCC CTAGGGGTCA AGGCTGCAGT GAGCCAAGGT CACGCCACTG CCCTCCAGCC TGGGTGACAG AGAGAAACAC TGTCTCAAAT 19000
TAGCCGGGCG TGGTGCGGCG CACGTGTAGT CCCAGCTATT CGGGAGGCGG AGGCAGGAGA ATCGCTTGAA CCCGGGAGGC AGAGGTTGCA GTGAGCCAAG 19100
ATAGCGCCAC TGCACTCCGG CCTGGGTGAC AGAGCGAGAC TCCATCTCAA AATAAATAAA TAAATAAATA TAAACTTTTT TTTTTTAATT TTAAGGCAGG 19200
```

MAGE-B Cluster Sequence



```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
GCTCTGGCTC NCGCTTGTAA TCCCNNCATT TTGGGAGGNC GANGCCAGCC TCATCGCTTA AGCACNGGAG TTCGAGGACT GGGAGGTGGA GGTTGTGGTA 19300
AGTCNAGATT GCGCGGCTNC ACTCCAGCCT GGAAGACAGA GAGAGCCCCT GTCACAAAAA AAAAAAAAAA AAAAAAAAAA AAGAAAGAAA GAAAGAAAGA 19400
AATATGAGGC CCACCGGTGC TAACCAGGGA NCCTCTCTCC AAAAGNNGGG CCCACCAAAA GCCCTAACCC TGTTTCAGGT CTTTGAAGCC CCAGGACATG 19500
GTCCGATNNC NTGCCTNGNC TTCCCCTCTA GGGACTNTGG GAGGGGAGGA TTTTGGAGGT TGGCGGACTT GGCTCNGTNG NGGTGTTACT CTGCTCCGCT 19600
TNGTATCNNG GTGAGNNACC TGNNTAAGGA CCTNGGGACC ACTGACTCCA GNACAGTGGG GTCCCAGCGT GTCNCCCGCT GCTGTCAGCC CTCGGAGACC 19700
CCGAGCGGGG TGTGGCTCNG CCTCACTTCC GCTTTGNNAG TGAGGCAGTT GGCTGNCGGG TGCNATNGCT TCAGTCAGGT TCGTGGCCTA GCGTGAGTCT 19800
TNGNACTGNT CCGGNGTNNA GGTNAGNNCC CTCNGCGGGC ACTGAAGGGA CNNTCCATGT TTTTNNCCCT GGNNAGCCCT NGGCAGNGCT ATNNNNCTGG 19900
CTTGCCTCTC ATTTCAGCTG GCGGGTCTCA GGGNGGGGAA GGCTTTCTCT NNCNGGGCNG CCTCNGTTCT TCNGNGGTCT CTTGGCCCTA ACTNGNGTCA 20000
NGATGNNGNC CNNGCCGGTA CGGCGTCTCA GGCCTGTNNT CCCNGCNCNC TGGGNGGCCG NGGCGNGTGN ACTGTTTGGG CCCCCCCNGG NGTTCGNNNC 20100
CNGCCGGGNC NNCNTGCGGN NNNCCTGTCT CTNCNNNNNN NNNNNNNNNN TGGCGGGGCN GCNGTGGNGC NCCCCTGTNG TCCCNGCTNC CCNGGNGGCT 20200
GNGGTGGGNG GNTCGCTTNN GCCCTNGCGG TCNNGGCTGC NGTGNGCCNN GGTCNCGCCN CTGCGCTCCN GCCTGGGTGN CNGNGNGNNN CNCTGTCTCN 20300
NNTTNGCCCG GCGTGGTGCG GCGCNCGTGT NGTCCCNGCT NTTCGGGNGG CGGNGGCNGG NGNNTCGCTT GNNCCCGGGN GGCNGNGGTT GCNGTGNGCC 20400
NNGNTTGCCC CNCTGCNCTC CCGCCTGGGT GNCNGNGCGN GNCTCCNTCT CNNNNTNNGT NNNTNNNTNN NTNNNTNTNN NCTTTTTTTN TTTTTTNTTT 20500
TNNGGCNGGG CTGTGGCTCN CGCTTGTNNT CCCNNCNTTT TGGGCGGNGG NNGCCNGCCT CNTCGCTTNN GCNCNGGNGT TCGCGGNCTG GGNGGTGNGC 20600
TGCCCCTGTN GTCCCNGCTN CCCNGGNGGC TGNGGTGGGN GGNTCGCTTN NGCCCTNGGN GTCNNGGCTG CNGTGNGCCN NGGTCNCGCC NCTGCGCTCC 20700
NGCCTGGGCG NCNGNGNGNG NCNCTGTCTC NNNTTNGCCG GGCGTGGTGG GGCNCNGGTG TNGTCCCNGC TNTTCGGGNG GTGGNGGCNG GNGNNTCCCT 20800
TGNNCCCGGG NGGCNGNGGT TGCNGTGNGC CNNGNTCGCG CCNCTGCNCT CCCGCCTGGG TGNCNGNGCG NGNCTCCNTC TCNNNNTNNN TNNNTNNNTN 20900
NNTNNNTNGN TNTNNNCTTT TTTTTTTTTT TNNTTTTNNG GCNGGGCTCT GGCTCNCGCT TGTNNTCCCN NCNTTTTGGG NGGNCGNNGC CNGCCTCNTC 21000
GCTTNNGCNC NGGNGTTCGN GGNCTGGGNG GTGGNGGTTG TGGTNNGTCN NGNTTGCGCG GCTNCNCTCC NGCCTGGNNG NCNGNGNGNG CCCCTGTCNC 21100
NNNNNNNNNN NNNNNNNNNN NNNNGNNNGN NNGNNNGNNN TNTGNGGCCC NCCCGTGCTN NCCNGGGNNC CTCTCTCCNN NNGNNGGGCC CNCCNNNNGC 21200
CCTNNCCCTG TTTCNGGTCT TTGNNGCCCC NGGNCNTGGT CCGNTNNCNT GCCTNGNCTT CCCCTCTNGG GNCTNTGGGN GGGGNGGNTT TTGGNGGTTG 21300
GCCGNCTTCG CTCNGTNGNG GTGTTNCTCT GCTCCGCTTN GTNNCNNGGT GNGNNCCCTG NNTNNGGNCC TNGGGNCCNC TGNCTCCNGN NCNGTGGGGT 21400
CCCNGCGTGT CNCCCCCTGC TGTCNGCCCT CGGNGNCCCN GNGCGGGNCG TGTCTCNCCC TCNCNNCTCC CTGCCCCTTN CNNNGGGGN CCCNCNCGTC 21500
TCNCCCTTGT GGTTGNCCCT GGNGGTTCCT GTTTGNGTTC TGTCTGGNNN GGCNCCCNGN GGGNGGGTCT TTCNCTNNGG GNGCNGCCCC CNGTTNTTCN 21600
GTTGGCCGGG NCCTGGNCCC TNNCTGGNGC CNNGGTGNNG TCTCCGNGTG CTNNNGGNTG GGNTCTGTTC CCNGTNGGGG CGTCNNCNGN NNGCNGCNGC 21700
NGCCCTNCTG GTTNGCNTCC TCCTGGTTNG CNCCNGTGGT CCTTGTCTTT TTTNTTTNTT TTTNTTTTTT TTCNGNCNNG GTCTCNCTGC TGCTNGCCNG 21800
TTTNGNNGGC TCCNGNCNNG TGTNGTCNTN NGCGGNGGCC CTGNCTTCCC CNTCCNGGGN GGGNNGTGTT CNGGGNGNTG NGGGTTTTGT TTGGNGGTTG 21900
GCGNNCTCNG GTCNGTNGNG GNNGNNNTTT CNGGCTGTCN TNNGNTNCCN NGGTGNNGNC CCCTGNNNTG NGNNTCTNGG GNCCNGCNNC TCCNNNNCNG 22000
TGNNGTCTCN TNGNGTTCCN CCCGTGTTGT CNGCCNTCNG NCCCCNGGNN GCTGTGNNCN GNTNNGGCTC TTCCTCNCTT CCTTGGNNGT GCTTTGNNGG 22100
GGNGGNTCTG GNGGCGNGGG GCNCGGGNTC TCTTCGGCNG NCGGTGNNTT CCTTGGNCTG TCTGGNGTCN NGGTCNGGNC CCTGNNTGTG CNTGNNNGGG 22200
NCCNCCNTCC CCCNNCCTGT NNCNNNGNGG GCCNCNCTNN NTCCTGCCCC GGNNGTCTTC CCTGGGNNTN CCTTGNNNGC TGTCTGNCNG NTNTCTNCCT 22300
GGNNNGTCTC NCGGNGGGNN GGGCCTTGGT CTNNGNNNGT NGCCCCNGTT CNGCNGGTGG NNGNGNGNCT TGGGTCCTNN CTGGNGTCNN GGTGNGGGCC 22400
```

MAGE-B Cluster Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
CTGAGTGCTA ATGAAGCAAT CTCTCTGCAA TAGAGGTGTC AACACAAACA GTGTCCTTGC ACTCATTTTG CAACCTCCAG GCAAAGGTAT TCATAGGTGA 22500
GGGGTCCCTG ACTTCCCTGT CTAGGGTCTT GTTCTGAGGC TGGAGGACTT AGGTTAATGG AGGGAAGTGT CCCACATCCT ACTAAGAGTC AAGTCAGAGA 22600
CCTGAGAGAA AACTAAAGGG AACACTCTCT CTAAAAAGTA GAGTCCCACA AATTGTGCCT CTTCTCTCAG CTCCAGGAAG CTTTGGAATA ACGTCAGCCC 22700
CCCTTACTGC CTAGAGAGTG CCAGAAAGGT TGAACTAATT ACATGCCCCA GTTCTGCAGA GGGAAGAGTG AGGAGACCCA GAGCCTTACA GGGGTCACAG 22800
TGAGGACCCT GAATGAAGAC TAGTGATACA ACTCCTCCCA CATAAAGAAG AGACAACAAA GATTCCCCCA CTCCCCACTT GCAGCGCTGA TGAGGCTGGG 22900
CATGGAGGTC AGGTGTATGT GAACTCACCT TCTTTTAATG GCATTAGGTT GCTGAGAGCC TTTATCTAAA GTGAGGGGCA TCAGAGCAGC AGAGGCAGCT 23000
GTCCTAGGTC CTATCTAGAG ACAAGATGAG GACTATGAAT GAGGTGTCAG GACTCCAATA AGCCCAGGAA AGAGTAGGAC TCCACAATGC TGTTAGCACC 23100
CAGTTCCTCC TGTCAGGGGT GGTAAGGCTG AGACATTCCT TCACCTCCTC TTAGGTGGTC CCAAGGAGAT GAGGACATTT GCTGGAGGTG TCAAATTTAG 23200
TGCAGCACAG GGGAGAAGAC CCAGCCCTGA CAGTTTTTAT GATGGTCCTG AGTGTGAACT GAGAGAAACC TCCTACCCCA GAGTAAAAGG AGATCCACAA 23300
GGACTAGACA TGCCACGGCT GCTCTCAGTC CCAGTATACA CAGAACAGGG CTGGCAGGCT GTGGCCTAAG GCACACTTTA ATTACTTTCA CAGGGTTCTC 23400
AGAGGACAAA CTGATCAGAA CAGAAGCCTC TGGGTTTCCA GAGCAGTGCT CTCACAGAAA ACTGCAGAGG CGACCTTCTT TTAAATCCAA AGTGGTACCT 23500
CTCTGCTGAA GGCACTCATA CCCTCTCTTT CTCTCTCTCC TCCAGGTGCC TGTATCACCT GCCCTTCTGC TGACACTCCT GCCTGCTGTT CCTGACTACA 23600
GCCATCATGC CTCGGGGTCA GAAGAGTACG CTCCATGCAC GTGAGAAACG CCAGCAGACC CGGGGTCAGA CCCAGGATCA CCAGGGTGCT CAGATCACTG 23700
CAACTAACAA GAAAAAAGTA TCCTTTTCAT CCCCTCTTAT TTTGGGGGCT ACTATCCAGA AAAAGTCTGC TGGTAGGTCA CGTAGTGCTC TCAAGAAGCC 23800
TCAGAGAGCA CTATCCACCA CTACATCTGT AGATGTTTCT TACAAAAAGT CATACAAGGG AGCCAACAGC AAAATTGAGA AAAAGCAAAG CTTCTCTCAG 23900
GGTCTATCCT CCACTGTGCA GTCTCACACA GACCCTCTAA CCATGAAGAC AAATATGTTG GTGCAGTTCC TGATGGAAAT GTACAAGATG AAAAAGCCCA 24000
TTATGAAAGC AGATATGCTA AAAATTGTCC AAAAAAGCCA TAAGAATTGC TTCCCTGAGA TCCTTAAAAA AGCTTCTTTC AACATGGAGG TGGTGTTTGG 24100
TGTTGATTTA AAGAAAGTTG ATTCTACCAA GGACTCCTAT GTCCTTGTCA GCAAAATGGA TCTCCCCAAC AATGGGACAG TGACTGGTGG GAGGGGATTT 24200
CCCAAGACAG GTCTCCTGCT GAATCTCCTG GGGTGATCT TCATGAAGGG CAACTGTGCC ACTGAGGAGA AGATCTGGGA ATTCCTGAAT AAGATGAGAA 24300
TATATGATGG GAAGAAACAC TTCATATTTG GGGAGCCCAG AAAGCTCATC ACCCAAGATT TGGTGAAGCT TAAATACCTG GAGTACCGAC AAGTGCCCAA 24400
CAGTAATCCT GCACGCTATG AATTCCTGTG GGGTCCAAGA GCCCATGCTG AAACCAGCAA GATGAAGGTC CTGGAGTTTT GGGCCAAGGT CAATAAAACT 24500
GTCCCCAGTG CGTTCCAGTT CTGGTATGAA GAGGCTTTGA GAGATGAGGA AGAAGAGTC CAAGCTGCAG CTATGCTCAA TGATGGCAGT AGTGCCATGG 24600
GCAGAAAGTG TTCCAAGGCC AAGGCTAGCA GCTCTTCCCA CGCCTAGTGA AGTTGAAGCA AATTTTGCAT TTTGTGGTTA AAGAGGGCAG TCACTGTTCC 24700
AAGGAGTGAA GGACTGGGTG TTACTGGAGG GAACACACTG TATAATACCT TTTGTTTCTG TTCTAAATGG ATAATTTGAA GTTTTATCTG TATTTTGGGG 24800
CATATTTTTC AAATGTTCCT TTTATTTAAC ATTGTAATCT AAGTTTAGGA TTGATACTGG TCACATTTGT TGTTTAAGAG TAAAAATTTT GCTGTTTTGT 24900
AAAACAGATT GAGAAAAATT CGATCTTATT TAGTGATCTG TTGCAAGATA ACTTGGAATT AGAATAAGCA TTTCCTTGAA AATGTTTAAA AAAAAAAAGT 25000
CAGCAGTAAA ATGTATGGCA TTAAGAAATA GAGAAAGAGT GTAAGATGGT CAATATTTGG TTTCCTAAAT GCTTTTACTC TGTGTTTTAA GAAAATGAAA 25100
GATAAATAAC CATATATGTC TGGCTTACTT AAGAATGTAG AATTAAATCA TAATAAATTA GACCTCATGC TGACACACTC ATTCTCCAAG TGTTAATTGA 25200
GCATCTGCTC TTGGAAGGAT CCATGCTAAT ACTGGGAGGG CTAAGAAAAA GAAGACCTAG CAACTGACCT TAAAATTATA AGGTGGAGAA GCAGCTATCA 25300
TCTAAGGAAA ATGGTGATAT ACACTCTAAG ACCAAAAGGA TACATGATAA GAAGGGAGGG AGGTGGTTCC ACATGAGAAC AGTCAAGTAT GAATTATCTA 25400
ATCAAGGGGG TGTTGGGCCT AAGGAAAGTG CAGGTCCCTC AATGGAGTT AATCTAAGTG AGGCTCCATG GTGGGCTGGA TGAGCCTGTG GGAATGGGGA 25500
GCAGGGGCCA GGCTCTTAGA AGTTGCCTCT CACGCAGTGC GGGCGCCTGT AGTCCCAGCT ACTGGGAGG CTGAGGCAGG AGAATGGCGT GAACCTGGGA 25600
```

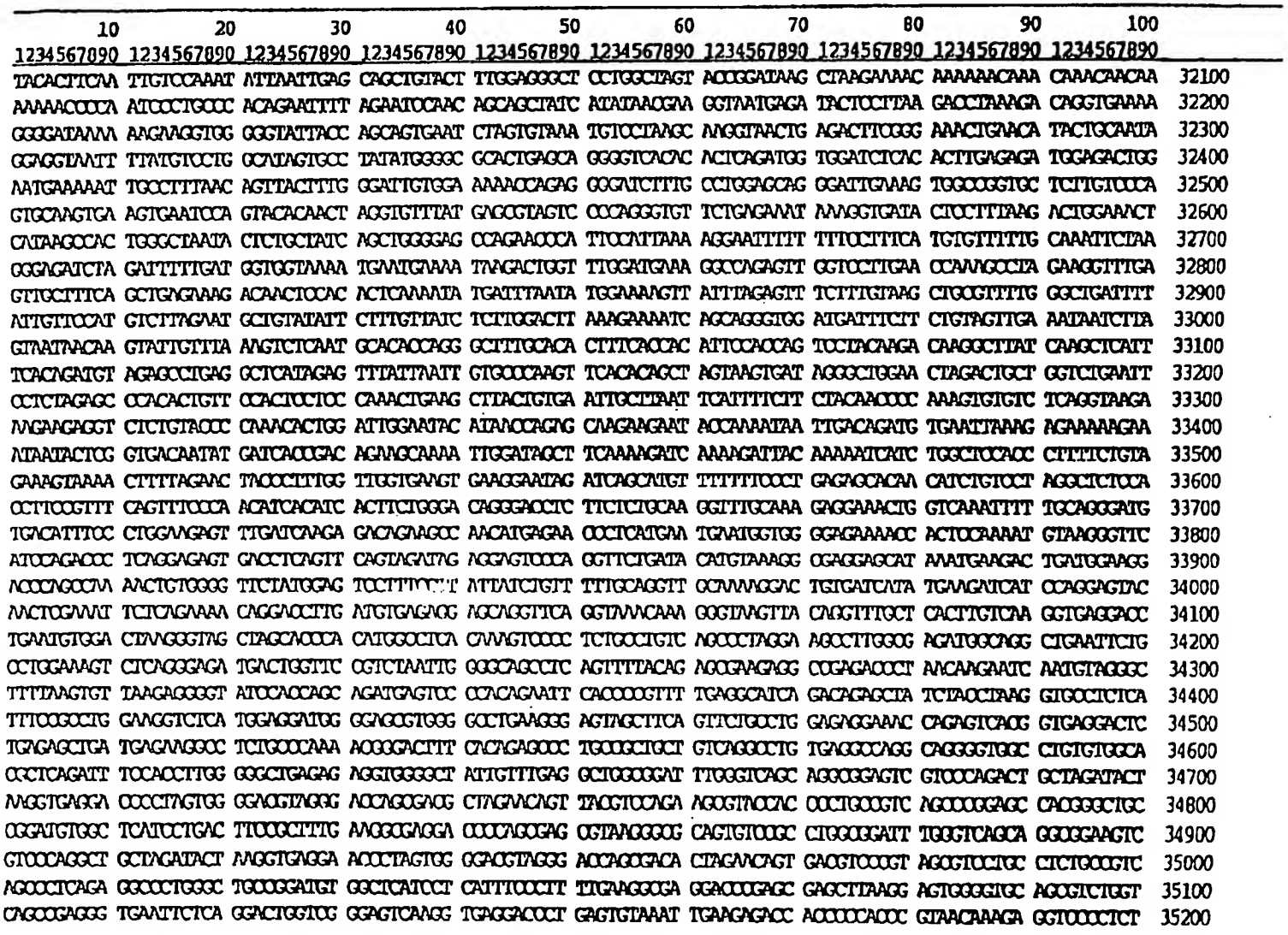
MAGE-B Cluster Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
GGCAGAGCTT GCAGTGAGCC GAGATTGTGC CACTGCAGTC CGGCCTGGGC TAAAGACGGG GATTCCGTCT CAAAAGAAAA AAAAAAAGAA GTTGCCTCTC 25700
AGAGGTGCGA GACAAAGCTT GGAATAGGAA ACAGTTCTTA ACAGTTACTT TGGGATCATG GATAAATCAG GGAGAATCTG GGGTAGGAAT GTGTCCTGTG 25800
CTCTTGTCCC AGTACAAATA AACACAGCAC ACACTAGGTG TTTTGTGCAC ATCATCTCCA GCCAGTTTCT GAGAAATAAG GGTAATACCC TCAAGGGACG 25900
GGGGAGGCCC AGAAGCCACT GTGCTGGTAC TCTTTGCCCT GGGCTGGGGG ATCCAGAGCC TGTTCCATTA AAAGCGCATT CAATTAGGTT ACCTCATTTG 26000
TGATTTGGCA GACTCTAGGC AAGTTCTAGA TTTTGAGAGG TGGTTAAATG AATCAAAATA AAAGTGGTTT AGATGGAAGG ATAGCTGAGA GAGAGGGAAG 26100
TTTTTGGTCC TTGAAACTCA TTTTAGAATC TGTGTTGCAT CTACTGAGGA AGTCTACCTC ATACTAATAT TGAATGATCT ATCAAGTAAT GAGGAAATAT 26200
TATTTAGTTT TCCTTTCTAG GTAGCACTTT GTCTGATAGG GCTTTCTTTG TCTTAGAATA TGGTATAATC TCTGACATCT CCTGCATTGA AAAATAAGTC 26300
CCAGTGTAGT AGATTGCAAC GTCAGGGTAC AAATTAATCA TAATAAGAAC TGGCACTATT ATATGTCTCA ATACATCTCA AGCACTTTAC ATACATTGTA 26400
TACATTTTAC CAGTCCTACA AAACAGAGCT TATCAAACCC ACTTCACAGA TAAAAAGCCT GAAGCCCACA GTCTATTAAT ATGTCCAATT TCACATGGCT 26500
GGTAAGTGAC AGGACTGGGA CAAGACCCCC AGTCTGAATT AGTGTAAAGC ACACTCTGTT CCATGCCTCC CAACCTGAGA CTGACTGTGT CCCCTTATTC 26600
ACTTTTCTTC TCACTACTCC CAAACGTATA TCAAGTGATG AAAAGATTCT GTACCCAATC TGCTGGACTG AAATACAGAA CCAGAGCAAG GAAATTACCA 26700
AAATGGTCAA GAGGTGTGAA TTAAGGAGAG AAGATAATAT TTGGTGACAA TAAGATTACA CATGCAAAGG CAGATAAACT CAAAAGTCAT GGGGTCACAT 26800
TCTGGCAATC TGGACTGGCC CCCCTTTTAT AGGAAGGTAA ATCCTTTCCA AGAGAAATAC ATAAGAGGGT AGTAAAGAGA AGAAATAGGA TCAATGGTTT 26900
TTTTTTTTTT TTTTTTTTCT CTGACAGTAT GACTTCTATC CTGGCTTCT GCCTTGTGTT ACAGCTTCCC TATCTTATAT CACTCCTGGG ACAGGGATGC 27000
ATTGTCTGCA AGATTTATTA TAGGGAAACT GATCAAGAAT TTTTAGGGAT GCGGCATTTC TCAGGAAGTC ATTAACTAAG GGACAAGAGC CAAGATAAAG 27100
ACCTTCATAT AGGAGGACTC AGCAGAACAA CATGCCAGAC AGAGTTCTTA GCCTGAGCAT CTCCTCACTT ATTCTTCAAG GTTCTCAGGG ACTTGTGAGT 27200
TCCGCATCTC AGGTCAGTAG AGGGAGGAGT GCCATTCTCT GACACATACC AAGTTGAGGA CCCTGAATGA AGAATGAGGG AAGCACTCAC CCAAATAGCC 27300
TTAACAGGAT GTGGCTGAAC CTGATTTCTA CTCCGGAGGT CTTAAGAAGG TAAAGACCAT GATCTGAGGC TGGCTGACTC AGGTCCATAG AGGAAGGATT 27400
CCTAATGTGT GCCAGGAGTA AAGTGAGTAC CCTGAAGAGT GCAGGAACCA CCAAACCCCT ATCAGTGGGA TCCCACAGAA TCCTCCCTAT TGTCACCTCT 27500
GAGAGATTTA GGCATGAATG TCAGAAAGAG GCACCCTCAT ATCCTCAGTA GTAACAGAGA AGTGAAGGTC TTGTTCTGAG GTGGGCAATC CCAGCTCAGC 27600
AAAGGGAGGA ATATTAAGCC TTCCTGTGAG TGAGATTTAG GATCCTGAGT TAGGACCTGA GTAGGACAAT CCACCATAAG ACCTCAACCT GTCTACCACA 27700
GCTGCCACTC TTGGGAGACC ATGGTCAGCT GTGCCAGATA AGGCCACCCT CTTCCTCCTA TCAGATCTCA GGGAGTGCAG CCTTTTTAAA TGGAGAGTCT 27800
TCAGTCAAAA AAGAGGGGAG CCCTAGGATC TGCTAGGAGC CAAATCAAGG AAACCGAGTG AGAACTGAGG GGACTGCTCA CCTCTGAATA GAGAGAGCAA 27900
CAGAGTCCAC CCCCCTGCAT AACTCTTAAA GTGCCAGGGC AACATGTAAG GCTGAGGACT CACCATCACT TTTTTTTTTT TAATAATTGA CCTTCTTTTT 28000
TAGAATAGTT TTACGTTCAC AGCAAAACTG AGCAGAAAGG AGAGTCTCAC AGGCATCCTG TCCTCCCACA GACACAGCCT CCTTCATATC AACCTCCCAA 28100
GCCACGGTGA TACATTTGTT ACAATTGAGA ATACATGGAC ACATCATTAT CACCCAATGT CCGTAGTTTA GGATTCATTC CTGTTGCTGC ACATTCTGTG 28200
GTTTTAATAT TTAAAAATGT ATAATGACAT ATGTTACCAC CATAGAATCA TCCAGAAATT GTTTTACTGC CTTAAAAGTC TCTGTGCCTG ATCTATTTAT 28300
CTCTCACTCT CCCCAACCTG TGGAAACCAC TGAACTTTGT ACTGTTTCCA TACTTTTGCC TTTTCCAGAA TGCATATAGA TGGAATTATG CAATGTATAG 28400
ACTTTTCAAA TTGGCCTCTT TCACTTTGTA ATATGTATCT AAGTTTCATC CATGTCTCTT TGTGGCTTGA TAGTTCATTC CCTTTTAGTG CTGAATAATA 28500
TTTTATTCTG TTAATTTATC CACTCTCCTT CTGAAGGACA TATTGCTTGC TTCCCAGTTT GGGCAATTAT GAATAAAGCT ACTATAAACA TTTGTGTGCA 28600
GGTTTTTTTT GTGGACATAC CGTTTTCAAC TCATCTGGCT AAATAACAAG GAGCGCACTT GCTAGATCAT ATGATGAAAG TATGCATAGC TTTGTAAGAA 28700
ACTTACAAAC TGTCTTCAAG AGTGGCTGTA TCATTTTTCA TGTCCACCAG CAATGACAGT GTGCTCCTGT TGCTCCACAT CCTCATCAGT ATTTGGGGTT 28800
```

MAGE-B Cluster Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
ATCAGGGTTT TTGATTTTGA CCATTCTTAT AGGTATGTAG TGGTATCACA TTATTGTTTT AATTTGCAAG TCTCTNATGA TATGTGATGT TGACCATCAT 28900
GTCATGTACG TATTTGCCTT CCATATTATAT TATTTGTTGA GGTGCCTGTT TAGATCTTTT GTCCATTTTT AATAGGGTTG GTCAGTTCTT ATTGTTCAGT 29000
TTTAAGACAT TTTTGTATGT TTTGGTTAAC AGTCCCTTAT CAGATATGTC TTTTGAAAAT AATTTTTTCC CAACCTGGGA GTTATCTTAT TCTCTTTGTG 29100
GTATCTTTAG CAGAGCAGAA GTTTTAATTT AGCGAAGTCC AGATTATCAA TTATTTTCTT TCATAGATGC CCATCACATT TTATACCAGA ACTAGAAAAG 29200
ATGAAATTCT TGGTCTGAAG TTGCAGTTAT CAGTCAGCAG AAGAGACAGT CCACAACCCT GCTTGGAGTC CAGATGAGGA TCCTGAGTGC AAACTTGGGA 29300
CCTAAAGAGC CCAGGACAGA GAGAGCACTA AATGCTTCTA GGCAGGGGTG GTGGGTTGAG GGGCCCCTAG ACTTCCCTCA TCTGGGTCCC AGAAAACTAA 29400
AGAGTCAATT TCACAACACC AATAGAGGGA GGCTCAGGCC CTGCCAAGAG CTGACATGAT AATTCTAAAG GTAATCAGAG TGGATCCTCT CCAAGCCAGA 29500
ACACAGAAAG CCCCACTGCG AGCCTTGTTG TCACCCAGTC AGCCCCAGGC AGGGTTGGCA AGCTGCAGCC TAAGGCACAT TGTAACTTCC TCAGCTGGCT 29600
TCTCAGGGGA CAGAATGACT AAGAACAATA GCCCAGTGAA TACTTAGAGC AGTGTTCTCA AGGAATCCTG CAGAGGCGGC TTCTGAAAAG CCAAGGTAGT 29700
ATCTGCCTGC TGAAGGTGTT CTCAGGATTT CATTTGCTCT TCTCCAGGAA CCACATCACC TGCCCTTCTG CCTACACTCC TGCCTGCTGT GCCTAACCAC 29800
AGCCAATCATG CCTCGGGGTC AGAAGAGTAA GCTCCGTGCC CGTGAGAAAC GCCAGCGGAC CCGTGGTCAG ACCCAGGATC TCAAGGTTGG TCAGCCTACT 29900
GCAGCAGAGA AAGAAGAGTC TCCTTCCTCT TCCTCATCTG TTTTGAGGGA TACTGCCTCC AGCTCCCTTG CTTTTGGCAT TCCCCAGGAG CCTCAGAGAG 30000
AGCCACCCAC CACCTCTGCT GCTGCAGCTA TGTCATGCAC TGGATCTGAT AAAGGCGACG AGAGCCAAGA TGAGGAAAAT GCAAGTTCCT CCCAGGCCTC 30100
AACATCCACT GAGAGATCAC TCAAAGATTC TCTAACCAGG AAGACGAAGA TGTTAGTGCA GTTCCTGCTG TACAAGTATA AAATGAAAGA GCCCACTACA 30200
AAGGCAGAAA TGCTGAAGAT CATCAGCAAA AAGTACAAGG AGCACTTCCC TGAGATCTTC AGGAAAGTCT CTCAGCGCAC GGAGCTGGTC TTTGGCCTTG 30300
CCTTGAAGGA GGTCAACCCC ACCACTCACT CCTACATCCT CGTCAGCATG CTAGGCCCCA ACGATGGAAA CCAGAGCAGT GCCTGGACCC TTCCAAGGAA 30400
TGGGCTTCTG ATGCCTCTAC TGAGTGTGAT CTTCTTAAAT GGCAACTGTG CCCGTGAAGA GGAAATCTGG GAATTCCTGA ATATGCTGGG GATCTATGAT 30500
GGAAAGAGGC ACCTTATCTT TGGGGAACCC CGAAAGCTCA TCACCCAAGA TCTGGTGCAG GAAAAATATC TGGAATACCA GCAGGTGCCC AACAGTGATC 30600
CCCCAGGCTA TCAATTCCTG TGGGGTCCAA GAGCTCATGC AGAAACCAGC AAGATGAAAG TCCTGGAGTT TTTGGCCAAG GTGAATGACA CCACCCCCAA 30700
TAACTTCCCA CTCCTTTATG AAGAGGCTTT GAGAGATGAA GAAGAGAGAG CTGGAGCCCG GCCCAGAGTT GCAGCCAGGC GTGGCACTAC AGCCATGACT 30800
AGTGGGTATT CCAGGGCCAC ATCCAGTAGC TCTTCCCAAC CCATGTGAGA TCTAAGGCAA ATTGTTCACT TTGTGGTTGA AAGACCTGCT GCTTTCTCTG 30900
TTCCTGTGAT GCATGAATAA CTCATTGATT TATCTCTTTG TTGTATTTTC CATGATGTTT CTTAAAATAG AAAGTTTATT TAGATTCAGA ATATAAATTT 31000
AGAAATGGCA TGCATCACAC ATTTATTGCT GTTTATCAGG TTGGTTTAGT GATAATAATT TTGTTTTTGA AATACAAATA GAAAATCCTG AAATAATTTT 31100
TGTGATACAG AGCAAAATAA CAGGGCATGG GAGTAAGGTT ATCCTTAGAA ATTTAAAATA ACTCCACAGT AAAATAGGTA GAATCTGAAG ATAGAAAGGG 31200
AAGAAAAGTA AAAGTTGCTT TATTGGTGGT TTGTCTTACT CAGTTCAGTC TTTTTTTGCT CATAAATTTA AAAGTTACAT ACCTGGTTTG CTTAGATTAT 31300
TCAAGAATGT GGAGGCCTGG GCCAAGGTCA ATGACAGTGT CTCCATTGTC TTCCCTCCAT TAAGAGAAGA CTTTAAGAGA TGAGGGAGAG AGAGCCAGAG 31400
ACAGTGTTGC AACTGGGCCT GGCATGTTTC AGTGTGGTGT CCAGCAGTGT CTCCCACTCC TTGTGAAGTC TGAGGTATAT TCTTTACTTT TGATTAAGAA 31500
AACACTTAAC CTTCTAATTA ATGGAGAGCC AAAGGGGAGT TGGTGGGAAC ACCATGTATA ACATATTTGT ATGTAAAATG ATTTATCTTT TCTTTTTCCT 31600
GTTTTTCAGT GTTCTTTTTT TAAATTGTAG ATTTATTTAG TTTCAGAATC TAAGTTTATG AATGGCATGA ATCACTCATT TATTAAAATA TATCAGGTTG 31700
GAGAGTGAGA ATTTTTGCAT TATGTAAAAC AATTTAAAAA TCTTTTAAGT CTTTTTCTGT GATCTAGAAC AAGATAATAT GGCATTGGAA TATGGAATTT 31800
GTGAAAAGGA AATTACCTTG CAATAAAGTT GGTGGGACCA GGAAGTAGAG AAAAAAAAAG TAAAATGTGG TCAATTCTTG CTTTGTTTTA TTCTTTTTAG 31900
TCTTTCTTTG TAAAACTGAA GTATATGTAC CCGGATCTGC TTAGCTTTTT CAGGAATGTG GGGGAAATTA AATAGCAATA CATTTGACTT CCTGGTCACT 32000
```

MAGE-B Cluster Sequence



MAGE-B Cluster Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
AAGTCCCGCT TCTGCATTTG GTCCTGGGAG GCCTCAGGTA ACCAGATGGG TAGCACCCTG ACTGTCTCTT CAGCGACTCA GGGAGACGAA GGCTTTGGCC 35300
TAAGCCTTAT AGACTCAGGT CAATAGAGGG AGGAGTCCTA AACCCTACTA CCCGTAAATC CAGAACTCTG GGAGGCCGAG GCAGGCGGAT CACGAGGTCA 35400
GGATATCAAG ACCATCCTGG CTAACACGGT GAAACCCCGT CTCTACTAAA AATAGAAAAA ATTAGCCAGG TGTGGTGGTG GCCTGCTGTA GTCCCAGCTA 35500
CTCAGGAGGC TGAGGCAGGA GAATGGCGTG AACCAGGGAA GCGGAGGTTA CAGTGAGCCG AGATTGCACC ACTGCACTCC AGCCTGGGTG ACAGAGCGAG 35600
ACTCAGTCTC AAGAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAGTCCCGC TCCTGCTGTC GGCAACAGCA GGCCCCCAGTC AGCCTTGGTG GGATGTGGCC 35700
CACTATGACT GTGAACTTAG GTCCAAGGAA TATGAGAACT TTTGTCTACG GGGCATGGTG TTAGGAGCAG TTGATGGGTG GAGTCCCAGA AAGAGTGCTG 35800
AGTGGAGGTG GAGACACTGA GTGAGGAAAT GCGGGCTACT CTATACAGGA GGAGACAAT TGCGCCTGAT GCTGTCCCTG GGAGTCCCAG GCCAGAGCTG 35900
GCAGGCTCAA GTCTCCCTGG CTTCTGCCTT CATGGTCTTA GAGAGAGGAG GGCCTTGGTT TAAGGCCTCA CTGCCCCAGT TCAGTAGAGA GATAGGAGTC 36000
ATACGCCAAG ACAGGTGTCA AGGTACAATT TCTTTATGAG GAATGAGGTA ACCGTCTAGT CCAGCAACAG GCAGGAACAC AGTCTTACCT TGGCGTTCAT 36100
TCCTTAGAAG CCAGGGTTAG ACTCTCAAGT TGAGATGCCC CTCATTTCCT TCAGGTACTG TTTCAGGTAC ATAATGGCTT TGGTTTAACA TGAAAGCCCT 36200
TGTTTAGTAC AGGGAGGAGG CCCAGTTCCT AAAAGATGGT GACACTGAGG GAGTATGAGG GGACCCCCTC CAAGAAAGAG GTAACACAGA TAGAGCCCTA 36300
TCCCCACTGA GACCTGGGAG ATCCAGGTAT GGTGACATGA TGAGTCTCAC TCACTTCTTC CTAGAGCATC TCAGGAAGTG GGAACTTCAT CAAAGGGGGC 36400
AGCCTTATGT TAGCAGCTAG AGATTCCTAG TAATTCCAGG AGTCAAAAGA AGACCCTGAG AACTGGGGAA CCACTCATCC CATAACAGTG AAAGAAACAC 36500
GGAATTCCTT CCTGATTTTC ATCCTTGGGA GATGATGGAC AATTGTAGCC AGATGGGAAA AGTTTCACTT CTTCCTCAGG GAGTTGTCGG GGGGTGCTTT 36600
GTATGACTGG ATAGGCGTTA GATCAAGACA GAGAAGTCAC TGAGGGAGGA ATGAGATGAG CATTTACCTA GAGAGTGGGC TTCACCAAGT TCTACCCCCT 36700
CTCAGCTCTG GGAAACCCCA GGCAGAAGTA CCCAGATGTG TCATCCCCTC AAGACTAGCC CTGGGTACTC AGAGAGGTGA AGGCTTTTGT ATGAGCCTGG 36800
AAGAATCAAG TGGGATTAGA GAGGAGCTCA GACTCTGCTT GAGTGAAGAC TAAAGAGAAC AACGACCCCA GAACAGTGGA GCCCCATAGA GGACGTGGTG 36900
ACTGGATGTG ATTCAGGCTT TCTTTCATCT TGGGACTATG AGGCAATGAG AATCTTAATC TGATGTAAGG GGCCTCAGGT CAGTAGAGAG AGGATTTCCA 37000
GGTTGTGCCA GGCCTCATAG AGAGGACTTG AGGGAACTCC CACCTCATCA GTGGGGATCC CTCAGAGTCC CTCTGTATGT CAGCGCTAGG AAGCCCCGAG 37100
CATAAATGTC AGAAATGCCC CTAAATTCCT CTTCAGAAGT AACAGGGAAA TGAAGGTCTT AGTCAGATGG GTTAGCAGGA GGGGTGGGAG GCATTTTAGG 37200
CCCTCCCAGG AGTCAAACTG GAGACCTTGA GTGAGGACCA TGAGTGGGGC TACACATCAC AGGGCTTCAA CCTGCCAGCC ACAGCTTTCA GCCTCTGGAG 37300
AATATGGGCA TTGTGACCAG ATACAGCCAC CCTCATTTCC TCCATTGGGT CTCAGGCAGA TGTAGGCCTT ACTATGAGTA AATGTCCTCA GGTGTAGAGA 37400
GAACAGAGCC CTAGGCCCTT CTGGGAGTCG AAGTGTGAGT GTGGACTGAA GGCACCATTC CCCCAAGCCC CCAACCCCCA CCCCAAATAG AGGAAAAACA 37500
ACGATGCTAG CCCTGTCTCT GCACTTAGCT CTGAAAGGCC TTGGCCAAGG GTTGCCAGGC TGAGACTTTA TTTCTTTGCA TCAGGTCTAA GGGAGGTGAC 37600
AGCTTTGGTC TGAAGATGCA GCACCAGTTA GCAGAAGACA GGTTCCCAGA ACTTAGATAT AGATGAGATG AGGACTCTGA ATTAAGATTG AGGTCCAACT 37700
AGCCCAGGAC AGAGAGAGTT CCATAGAACT GTCAGCACTG CCATCCCGCC AGCCCCCGGT AAGGATGGTA GGTTGAAGCA GTGCCTCATT TTTCTTTGTG 37800
GATTCCAGGG AGCTTTGAAG TGTCAGCTTC AGAGCAGCAC AGGAAGGAGT CCCAGACCCT TCCAAGAGTA GATATGAAGA TCCTGTATAT GAATTGAGAG 37900
GCCTTGAACA CAGAGGAGTC TACACTGCCA ACCTCTGCTG TCACCCAGTC AGCCCAGGCA GGTTTGGCAA CAAGAACCAG TGGTTCCTAG AGCAATGCCC 38000
TCAAGAAAAC CAGCAGAAGT GCTCTCTAAA AGCCAAGTTG TACCTCCCTG CTGCAAGTAC TCACAGATCT CATTCTCTCT CCTTCAGGTG CCACATCTCC 38100
TGCCTTTCTG CTCACTTTCC TGCCTGTTTT GCCTGACCAC ACCCATCATG CCTCGGGGTC AGAAGAGTAA GCTCCGTGCT CGTGAGAAAC GCCGCAAGGC 38200
GCGAGAGGAG ACCCAGGGTC TCAAGGTTCG TCACGCCACT GCAGCAGAGA AAGAGGAGTG CCCCTCCTCC TCTCCTGTTT TAGGGGATAC TCCCACAAGC 38300
TCCCCTGCTG CTGGCATTCC CCAGAAGCCT CAGGGAGCTC CACCCACCAC CACTGCTGCT GCAGCTGTGT CATGTACCGA ATCTGACGAA GGTGCCAAAT 38400
```

MAGE-B Cluster Sequence

```
          10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
GCCAAGGTGA GGAAAATGCA AGTTTCTCCC AGGCCACAAC ATCCACTGAG AGCTCAGTCA AAGATCCTGT AGCCTGGGAG GCAGGAATGC TGATGCACTT 38500
CATTCTACGT AAGTATAAAA TGAGAGAGCC CATTATGAAG GCAGATATGC TGAAGGTTGT TGATGAAAAG TACAAGGATC ACTTCACTGA GATCCTCAAT 38600
GGAGCCTCTC GCCGCTTGGA GCTCGTCTTT GGCCTTGATT TGAAGGAAGA CAACCCTAGT AGCCACACCT ACACCCTCGT CAGTAAGCTA AACCTCACCA 38700
ATGATGGAAA CCTGAGCAAT GATTGGGACT TTCCCAGGAA TGGGCTTCTG ATGCCTCTCC TGGGTGTGAT CTTCTTAAAG GGCAACTCTG CCACCGAGGA 38800
AGAGATCTGG AAATTCATGA ATGTGTTGGG AGCCTATGAT GGAGAGGAGC ACTTAATCTA TGGGGAACCC CGTAAGTTCA TCACCCAAGA TCTGGTGCAG 38900
GAAAAATATC TGAAGTACGA GCAGGTGCCC AACAGTGATC CCCCAGCCTA TCAATTCCTA TGGGGTCCGA GAGCCTATGC TGAAACCACC AAGATGAAAG 39000
TCCTCGAGTT TTTGGCCAAG ATGAATGGTG CCACTCCCCG TGACTTCCCA TCCCATTATG AAGAGGCTTT GAGAGATGAG GAAGAGAGAG CCCAAGTCCG 39100
ATCCAGTGTT AGAGCCAGGC GTCGCACTAC TGCCAGGACT TTTAGAGCGC GTTCTAGAGC CCCATTCAGC AGGTCCTCCC ACCCCATGTG AGAACTCAGG 39200
CAGATTGTTC ACTTTGTTTT TGTGGCAAGA TGCCAACCTT TTGAAGTAGT GAGCAGCCAA GATATGGCTA GAGAGATCAT CATATATATC TCCTTTGTGT 39300
TCCTGTTAAA CATTAGTATC TTTCAAGTGT TTTTCTTTTA ATAGAATGTT TATTTAGAGT TGGGATCTAT GTCTATGAGC GACATGGATC ACACATTTAT 39400
TGGTGCTGCC AGCTTTAAGC ATAAGAGTTT TGATATTCTA TATTTTTCAA ATCCTTGAAT CTTTTTTGGG TTGAAGAAGA AGAAAGCATA GCTTTAGAAT 39500
AGAGATTTTC TCAGNAATGT GTGAAAGAAC CTCACACAAC ATAATTGGAG TCTTAAAATA GAGGAAGAGT AAGCAAAGCA TGTCAAGTTT TTGTTTTCTG 39600
CATTCAGTTT TGTTTTTGTA AAATCCAAAG ATACATACCT GGTTGTTTTT AGCCTTTTCA AGAATGCAGA TAAAATAAAT AGTAATAAAT TATATTACTT 39700
GTTCAGTGCC TCATTTATTC TCACCATAAA TTGAGCATCT GCTCTTTGTA AGGCTCTGTG ATAGTAGTGA TTGTACTAAG TTAAAGAAGA CCCTTCGCCT 39800
GCACACAGAT TTTTAGTCTA AGGACAGTTA TTATTTAAAG AAGATGGTGA GATACACTCT AACATGTACA GATTTTTTTT TTTACATATA AACACTCATT 39900
TAAAAAAAAA AGAAGTGAGA ATGGTGGGAG AAGGTTCAGA CAAGAGCAGT CAAGTGTTAA TTTCCTAGCC AAGGCACTTC GTGGTGTGGG ACAATGCAAG 40000
TCCCTCGTTG GGAGGTCATT TTAAGTTAGC TCCATGGTGA ACTGGATGAG GTTGTGATAA TCATAAGAAG GTGCCAAACC CTCAGATCAT GAGCCTTAGA 40100
GTTGAGAGAT TAAGCCTGGA AAGGGAAACT GCCCTTAACA GTTACTTTGG GATTGTGGGT AAAGCAGAGA GAACCTGTGT CTGGAGGAGG AGTGGAAGAA 40200
TACAGTGCTC TCGTCCCAGG GCAGTCAAAC ACAGTGCACA AACTAGTTGT TTTATGCACA TTGTCTCCAG AAAGTGTTTG AGAAATAAGG GTTATACTTC 40300
CTTGAGGTGA GATGCCAAGA AGCCACTAAG CTAACACTGT TTCCCTAAGC TT                                                    40352
```

Claims

1. Method for screening for possibility of a testicular seminoma, non-small cell lung carcinoma, melanoma, breast cancer, sarcoma or leukemia in a sample, comprising contacting said sample with at least one nucleic acid molecule which hybridizes to mRNA corresponding to an MAGE-Xp gene, and determining hybridization as a determination of possible presence of testicular seminoma, non-small cell lung carcinoma, melanoma, breast cancer, sarcoma or leukemia in said sample.

2. The method of claim 1, comprising polymerase chain reaction.

3. The method of claim 2, comprising contacting said sample with at least two nucleic acid primers.

4. The method of claim 1, wherein said MAGE-Xp gene is MAGE-Xp2, MAGE-Xp3 or MAGE-Xp4.

5. The method of claim 4, comprising contacting said sample with (a) SEQ ID NO: 5 and SEQ ID NO: 6, (b) SEQ ID NO: 7 and SEQ ID NO: 8, or (c) SEQ ID NO: 9 and SEQ ID NO: 10.

6. The method of claim 5, comprising contacting said sample with SEQ ID NO: 5 and SEQ ID NO: 6.

7. Isolated nucleic acid molecule consisting of genomic DNA encoding a MAGE-B gene, consisting of nucleotides 3266-7979 of SEQ ID NO:15, nucleotides 23546-25194 of SEQ ID NO:15, 29748-31474 of SEQ ID NO:15, nucleotides 29748-31827 of SEQ ID NO:15, or nucleotides 31403-39691 of SEQ ID NO:15.

8. Isolated nucleic acid molecule which encodes a MAGE-B1 variant, consisting of, in 5′ to 3′ order, nucleotides 31403-31474, 33958-34062, 35057-35139 and 38088-39691 of SEQ ID NO:15; nucleotides 31403-31474, 33958-34062, and 38088-39691 of SEQ ID NO:15; nucleotides 35057-35139 and 38088-39691 of SEQ ID NO:15; and nucleotides 33958-34062 and 38088-39691 of SEQ ID NO:15.

9. Method for determining presence of skin carcinoma

mammary carcinoma, comprising assaying a sample for presence of mRNA for MAGE-B1 gene, wherein presence of said mRNA is indicative of possibility of skin carcinoma or mammary carcinoma in said sample.

10. Method for determining presence of leukemia, lymphoma, head and neck squamous cell carcinoma, or mammary carcinoma in a sample, comprising assaying said sample for mRNA for MAGE-B2, wherein presence of said mRNA is indicative of possibility of leukemia, lymphoma, head and neck squamous cell carcinoma, or mammary carcinoma in said sample.

11. The method of claim 9, comprising determining said mRNA by means of polymerase chain reaction which comprises using one of the oligonucleotides set forth in SEQ ID NO:11, 16 or 17 and the oligonucleotide set forth in SEQ ID NO:12 as primers.

12. The method of claim 10, comprising determining said mRNA by means of a polymerase chain reaction which comprises using the oligonucleotides set forth in SEQ ID NO:5 and SEQ ID NO:6, or SEQ ID NO:18 and SEQ ID NO:6 as primers.

13. Isolated nucleic acid molecule useful as an oligonucleotide primer, said isolated nucleic acid molecule having a nucleotide sequence as set forth in

SEQ ID NO:12,
SEQ ID NO:16,
SEQ ID NO:17, or
SEQ ID NO:18.

14. Kit useful in amplifying a MAGE-B gene, comprising a pair of oligonucleotide primers, said pair being selected from the group consisting of: (a) SEQ ID NO:12 and one of SEQ ID NO:11, 16 or 17, or (b) SEQ ID NO:6 and SEQ ID NO:18.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/09774

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C12Q 1/68; C12P 19/34; C07H 21/02, 21/04
US CL :435/6, 91.2; 536/23.1, 24.33
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2; 536/23.1, 24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	US 5,587,289 A (LURQUIN et al.) 24 December 1996, see especially columns 3, 4, and sequence ID numbers 3, 4, and 5-8.	1-7, 10, 12
A, P	US 5,612,201 A (DE PLAEN et al.) 18 March 1997.	1-14
A	DE BACKER et al. Structure, Chromosomal Location, and Expression Pattern of Three Mouse Genes Homologous to the Human MAGE Genes. Genomics, 1995. Vol. 28. Pages 74-83.	1-14

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 AUGUST 1997	12 SEP 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer AMY ATZEL, Ph.D. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/09774

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, CAPLUS, CANCERLIT, EMBASE, BIOSIS, INPADOC, MEDLINE, NUCLEIC ACID SEQUENCE DATABASES
search terms: MAGE, MAGE-B, MAGE-Xp, testicular seminoma, carcinoma, lung, hybridize, nucleic, skin, mammary, breast, leukemia, lymphoma, head, neck, squamous, polymerase chain, kit, primer